

Computational Heterogeneous
Electrochemistry – From
Quantum Mechanics to Machine
Learning

Thesis by
Yufeng Huang

In Partial Fulfillment of the Requirements for
the degree of
Chemical Engineering

The logo for the California Institute of Technology (Caltech), featuring the word "Caltech" in a bold, orange, sans-serif font.

CALIFORNIA INSTITUTE OF TECHNOLOGY
Pasadena, California

2019
(Defended December 4, 2018)

© 2018

Yufeng Huang

ORCID: 0000-0002-0373-2210

ACKNOWLEDGEMENTS

First, I would like to thank my PhD advisor, Prof. William A. Goddard III, for all the guidance and insight throughout my graduate years at Caltech. Bill is probably the most knowledgeable and most insightful scientist I can think of. The depth and breadth of the sciences being done by Bill and the Goddard group still amaze me to this day. I truly appreciate the opportunity to learn not just about sciences, but also the historical contexts of these sciences from Bill. When I started my graduate school, I set out to learn about the microscopic world, from quantum mechanics to chemistry to molecular dynamics to mesoscale observables. I am fortunate to practice research in such a wide range of domains with Bill's insights.

Next, I want to thank my thesis committee, Prof. Zhen-Gang Wang, Prof. Tom Miller, and Prof. Mark Davis for the exceptional teaching of the fundamental concepts in chemical engineering. The rigorous course works in thermodynamics, statistical mechanics, and reaction engineering form the basis of my approaches to understanding and solving heterogeneous catalytic problems in my research.

I would also like to thank Dr. Robert Nielsen (Smith!) and Dr. Tao Cheng for the all the advice and suggestions in many of the projects we worked on together. I still remember the early days of my PhD when Smith patiently taught me the nuts and bolts of quantum chemistry. Both Smith and Tao are extremely resourceful, and are truly the ones I could rely on whenever I encountered difficulties in research.

I also want to thank Dr. Lin-Wang Wang for hosting my short visit at the Lawrence Berkeley National Laboratory (LBL). Dr. Wang is an exceptionally insightful physicist and an amazing software architect. During the few months at LBL, I was lucky to work on Dr. Wang's new energy partitioning method and to implement machine learning approaches to solve chemical problems. The four months at LBL were some of my most productive months in my life.

Prior to visit LBL, I was invited by Prof. Haibin Su, now at the Hong Kong University of Science and Technology, to perform research in Singapore. I appreciate Prof. Su's kind hospitality and Dr. Qiang Wang's help during my short stay in Singapore. I also want to thank Nadine, Tara, Casey, Katharine, Adam, Richard, and David for sharing many great moments in Singapore and Southeast Asia.

In addition to the theorists, I also want to thank many of the experimental collaborators with whom we have worked closely. I want to thank Prof. Linyou Cao and Dr. Guoqing Li from North Carolina State University for working with us since the early years of my PhD. I also want to thank Prof. Haiqing Zhou and Prof. Zhifeng Ren at the University of Houston to engineer and understand some of the highest performing catalysts for hydrogen evolution reactions.

I would also like to thank the heterogeneous catalysis group (Hetcat) led by Prof. Manuel Soriaga (Manny) at the Joint Center for Artificial Photosynthesis (JCAP) on campus. During the early years, I was fortunate to be invited to many of the Hetcat discussions, which helped me enormously in understanding the experimental side of electrochemistry and surface science. Manny is friendly and welcomes discussions. In fact, most of my PhD work is motivated by the discussions with Manny during many of my random afternoon visits to his office. I also want to thank the other Hetcat members for sharing research ideas and discussions: Dr. Youn-Geun Kim, Dr. Alnald Javier, Dr. Jack Baricuatro, Dr. Kyle Cummins, and Dr. Chu Tang.

I also want to thank other JCAP experimentalists with whom I have collaborated: Dr. Jimmy Jones, Prof. Jesus Velazquez (now at the University of California Davis), Dr. Joshua Wiensche, and Dr. Yi-Rung Lin.

I would like to thank my former 07 officemates and now great friends, Dr. Samantha Johnson, Dr. Jason Crowley, Dr. Sijia Dong, and Dr. Robert Nielsen. I am really thankful to have met Sam during my PhD. In the office, Sam and I discussed science, philosophy, and many other ideas that truly expanded my knowledge. Outside of the office, Sam introduced to many activities, including volleyball, softball, and cycling, which have

become my most favorite sports since. My PhD would not be as well-rounded without having Sam as a great friend who I could discuss everything with. I also want to thank Jason for sharing many of the happy moments in the office. Jason has really sharp insights about everything that always allowed me to learn to look at things in new ways from our discussions. In addition, Jason's passion for cycling has really motivated me to become a better cyclist. Lastly, I want to thank Sijia for all the discussions about many subjects, many people, and many areas of science.

I would also like to thank my recent 07 officemates and also friends, Jin Qian, Yalu Chen, Dr. Soonho Kwon, and Dr. Hyeyoung Shin. We have shared many lunches, dinners, and office trips together. I want to thank Jin for sharing and exploring many research ideas from our respective catalysis projects. I want to thank Yalu for working and improving our new machine learning research direction, for being a cheerful cycling and running buddy, and for being a good friend to share many stories in our lives. I want to thank Soonho for the many activities and trips outside of the office. I want to thank Hyeyoung for enduring my endless conversations, for practicing badminton together, and traveling to many places.

I want to thank the other WAG members, including Dr. Caitlin Scott, Dr. Daniel Brooks, Dr. Saber naserifar, Prof. Qi An, Prof. Yuanyue Liu, and Prof. Ravishankar Sundararaman.

Moreover, I want to thank many friends and classmates for sharing my PhD experience: Dr. Kevin Shen, Dr. Kevin Fiedler, Dr. Kevin Barkett, Dr. Kevin Marshall, Dr. Mark Harfouche, Dr. Travis Schlappi, Dr. Jacob Hedelius, Dr. Karthik Ramachandran, Dr. Cristofer Flowers, Dr. Cheng Li, Dr. Hao Chu, Dr. Jin Yang, Weiyi Chen (USC), and Shengan Zhan (UCLA).

I would also want to take the chance to thank my college research advisor, Prof. Marc-Olivier Coppens, now at University College London, for motivating me to become a scientist. In the early days of my college study, Prof. Coppens's enthusiasm and passion really inspired me to become to scientist. My scientific journey would not have been possible without Prof. Coppens's nurturing guidance in those early days.

Last, I want to thank my family for all the support. I want to thank my mother for all the hard work to raise me and for allowing me to make the important decisions in my life to pursue my dreams. I also want to thank my late father, who taught me to be diligent and persistent, and who really invested in my education. I want to thank my brother and his family for helping me in any way possible.

ABSTRACT

Because of coulomb interactions and complex surface morphologies, rigorous methods for heterogeneous electrochemical catalysis were not well-established. Thus, for different types of electrochemical systems, a specific strategy must be adapted. In this thesis, we first used the cluster model to study the chemistry on a 1D chain of MoS₂ edges. Then, a rigorous grand canonical potential kinetics (GCP-K) method was developed for general crystalline systems. Starting from quantum mechanical calculations, the method gave rise to a different picture from the traditional description given by the Butler-Volmer kinetics. Next, we studied the chemical selectivity of CO₂ reduction on polycrystalline copper nanoparticles. Because of the complexity of the reaction sites, we combined the reactive force field, density functional theory, and machine learning method to predict the reactive sites on 20,000 sites on a roughly 200,000-atom nanoparticle. Such a strategy opens up new way to understand chemistries on a much wider range of complex structures that were impossible to study theoretically. Lastly, we formulated a machine learning force field strategy using atomic energies for amorphous systems. We have shown that such a method can be used to reproduce quantum mechanical accuracies for molecular dynamics. This method will enable the accurate study of the dynamics of heterogeneous systems during electrochemical reactions. In summary, we have developed quantum chemical methods and machine learning strategies to reformulate rigorous ways to study a wide range of heterogeneous electrochemical catalysts.

PUBLISHED CONTENT AND CONTRIBUTIONS

Chapter II : **Huang, Y.** ; Nielsen, R. J. ; Goddard, W. A. III ; Soriaga, M. P.(2015). “The reaction mechanism with free energy barriers for electrochemical dihydrogen evolution on MoS₂”. In: *J. Am. Chem. Soc.* 137, 20, 6692-6698. doi: 10.1021/jacs.5b03329.

This is an open access article published under an ACS AuthorChoice License, which permits copying and redistribution of the article or any adaptations for non-commercial purposes.

Huang Y. performed the calculations, analyzed the results, prepared the data, and wrote the manuscript.

Chapter III : **Huang, Y.** ; Nielsen, R. J. ; Goddard, W. A. III. (2018). “Reaction mechanism for the hydrogen evolution reaction on the basal plane sulfur vacancy site of MoS₂ using grand canonical potential kinetics”. In: *J. Am. Chem. Soc.* 140, 48, 16773-16782. doi: 10.1021/jacs.8b10016.

Adapted with permission from *J. Am. Chem. Soc.* 2018, 140, 48, 16773-16782. Copyright 2018 American Chemical Society.

Huang Y. came up with the theory, performed the calculations, analyzed the results, prepared the data, and wrote the manuscript.

Chapter IV: **Huang, Y.** ; Chen, Y.; Cheng, T. ; Wang, L.-W. ; Goddard, W. A. III. (2018). “Identificaion of the selective sites for electrochemical reduction of CO to C₂+ products on copper nanoparticles by combining reactive force field, density functional theory, and machine learning”. In: *ACS Energy Letters.* 3, 2983-2988. doi: 10.1021/acsnenergylett.8b01933.

Adapted with permission from *ACS Energy Lett.* 2018, 3, 12, 2983-2988. Copyright 2018 American Chemical Society.

Huang Y. came up with the concept, implemented the machine learning algorithm, analyzed the results, and wrote the manuscript.

Chapter V: **Huang, Y.** ;Kang, J.; Goddard, W. A. III ; Wang, L.-W. (2019). “DFT based neural network force field from energy decompositions”. In: *Phys. Rev. B* 99, 064103. doi: 10.1103/PhysRevB.99.064103

Reprinted with permission from *Phys. Rev. B.* 2019, 140, 48, 16773-16782. Copyright 2019 by the American Physical Society.

Huang Y. implemented the machine learning algorithm, molecular dynamics simulation, and heat conductivity calculation, analyzed the results, and wrote the manuscript.

TABLE OF CONTENTS

Acknowledgements.....	iii
Abstract	vii
Published Content and Contributions.....	viii
Table of Contents.....	ix
Chapter I: Introduction	1
1. Background.....	1
2. Quantum chemical approaches	2
3. Overview of the thesis.....	4
References.....	8
Chapter II: The Reaction Mechanism with Free Energy Barriers for Electrochemical Dihydrogen Evolution on MoS ₂	13
1. Introduction	14
2. Computational methods	15
3. Model systems	17
4. Results and discussions	20
5. Conclusion	28
References	29
Supporting information	33
Chapter III: The Reaction Mechanism for the Hydrogen Evolution Reaction on the Basal Plane Sulfur Vacancy Site of MoS ₂ Using Grand Canonical Potential Kinetics	34
1. Introduction	35
2. The grand canonical potential (GCP) formulation using the constant charge condition.....	38
3. Simulation Model for the basal plane of MoS ₂	41
4. Relationship between grand canonical potential reaction kinetics and Butler-Volmer reaction kinetics.....	42
5. Discussion of the HER mechanism on MoS ₂ basal plane.....	48
6. Overall kinetics	52
7. Comparison with other Group VI transition metal dichalcogenides ...	55
8. Conclusion.....	58
References.....	58
Supporting Information	64

Chapter IV: Identification of the Selective Sites for Electrochemical Reduction of CO to C2+ Products on Copper Nanoparticles by Combining Reactive Force Field, Density Functional Theory, and Machine Learning	68
1. Introduction	68
2. Machine learning and CO adsorption	69
3. C2 coupling	72
4. Conclusion	77
References.....	77
Supporting Information	81
 Chapter V: DFT based Neural Network Force Fields from Energy Decompositions	 90
1. Introduction	90
2. Methods	93
3. Results and discussion.....	102
4. Conclusion.....	110
References	110
Supporting information	116

Chapter 1

INTRODUCTION

1. Background

Energy is essential for human activities. The ability to control and generate energy contributes to the development of sophisticated human society¹⁻². As the average quality of life is improving throughout the world and the total population is on the rise, the demand for energy is projected to increase accordingly³⁻⁵. Since the industrial revolution, fossil fuels including coal, petroleum, and natural gases have been used to generate energy^{1,6}. These sources were the result of natural photosynthesis over millions of years⁷⁻⁸. Without adequate replenishment, the ever-increasing usage of limited natural resources is unsustainable. Therefore, to meet energy needs in the long term, energy must be harvested directly from much larger energy sources⁹. Thermal¹⁰, biological¹¹ and photoelectrochemical¹² approaches have been suggested and employed to harvest solar energy. In addition to harvesting energy directly from the sun, other renewable sources like hydropower, geothermal power, and wind power can be exploited¹³. Because all of these processes generate electricity, developing efficient electrochemical processes to store the renewable energy in chemical forms is universally applicable and can be easily integrated into these systems.

The application of electrochemistry in energy includes batteries¹⁴, fuel cells¹⁵, supercapacitors¹⁶, and electrochemical chemical productions¹⁷. Because of the abundance of H₂O and CO₂ on Earth, they are promising for large-scale energy conversion into chemical fuels in the form of H₂¹⁸ or carbon based fuels¹⁹. To perform efficient chemical transformations, catalysts are used to lower the energies of the reaction intermediates and reaction barriers.

Developing an efficient electrocatalyst for water splitting and CO₂ reduction to chemical fuels requires an atomistic understanding of the catalytic site. There are two main classes of catalysts, molecular or homogeneous catalysts and heterogeneous catalysts. Homogeneous

catalysts are isolated molecules dissolved in solution and typically involve metal centers. They are easier to study because the intermediates can be transferred and isolated for analysis. However, in electrochemistry, electrons tunnel from or to the electrode to drive reactions, and only the molecular catalysts near the electrode are directly electrochemically active²⁰⁻²², resulting in ineffective use of molecular catalysts in electrochemical applications.

On the other hand, heterogeneous catalysts are in a different phase as the reactants, making them easy to handle and suitable for large-scale industrial applications. In addition, heterogeneous catalysts can be assembled in different scales for further optimization²³. However, because catalytic sites on heterogeneous catalysts cannot be isolated for chemical analysis, surface science techniques are generally used to study these catalytic sites²⁴. Due to the limitations of current experimental techniques in determining the active sites and reaction intermediates during the reactions, theoretical insights are necessary to understand the underlying electrochemical processes.

2. Quantum Chemical Approach

Chemical properties are governed by the Schrodinger equation. Because the motions of electrons follow the Schrodinger equation, the solution of this equation describes bond-breaking and formation in molecules. However, an exact solution to the Schrodinger equation is computationally prohibitive for large systems, because for N electrons in $2K$ one-electron spin orbitals, there are $(2K)!/(2K-N)!N!$ Slater determinants in the full configuration interaction solution²⁵⁻²⁶. Thus, for most systems, the Schrodinger equation can only be solved approximately²⁷⁻²⁸. The Hartree-Fock method is the simplest approximation to the Schrodinger equation based on the mean field theory. In this approach, electrons interact with each other via an average potential such that the overall wave function is the antisymmetrized product of the orthogonalized eigenstates. To obtain chemical accuracies, post-Hartree Fock methods can be employed.

Another method similar to the Hartree-Fock approach is the Density Functional Theory (DFT), which approximates the Schrodinger equation using an energy functional based on

the total electronic density. In the original formulation, Hohenberg and Kohn established the uniqueness of the ground state electron density in the presence of external potentials²⁹. One of the simplest form of such orbital free DFT approach is the Thomas-Fermi functional based on the Fermi gas of metallic systems³⁰. However, such an approximation still performs poorly for general systems. An important breakthrough was made by Kohn and Sham, who re-formulated the mathematical proof using electronic orbitals rather than electronic densities³¹. Practical functionals including the local density approximation (LDA), PBE³², and B3LYP³³ functionals have been used to study a wide range of chemical systems.

The application of quantum mechanical calculations to catalysis typically involves calculating the relative energies of reaction intermediates and the transition states between these intermediates. Because a catalyst will be regenerated after each reaction, the overall catalytic process consists of a catalytic cycle, in which the catalyst interacts with the reactants, performs catalytic transformation, and then releases the products to regenerate itself for the next cycle.

For reactions in solution, temperatures and pressures are typically held constant. Thus, the appropriate thermodynamic potential is the Gibbs free energy³⁴:

$$G = U - TS + pV = E + ZPE + H_{\text{vib}} - TS_{\text{vib}} + G_{\text{solv}}$$

Where E is the electronic energy calculated from the Schrodinger equation, ZPE is the quantum mechanical zero point energy, H_{vib} and S_{vib} are thermodynamic contributions due to the vibrational modes, and G_{solv} is the solvation energy³⁵.

For homogeneous catalysis, experimental methods can be used to isolate and identify stable reaction intermediates. By adding electrons and protons in separate steps, the corresponding reduction potential and pKa for each intermediate can be obtained. In theoretical calculations, because homogeneous catalysts are isolated molecules, net charge can be placed on these molecules, which also allows the addition or removal of electrons and protons. Thus, the intermediate species can be calculated and compared with experimental observations.

On the other hand, heterogeneous catalysts are relatively more complicated because each reactive site is small compared to the overall structure. Perfect crystalline heterogeneous catalysts can be represented by a simple unit cell, which can be repeated to obtain the actual large scale structure. However, unlike molecular calculations, net charge cannot be applied directly to the periodic calculation because charge interactions between periodic cells result in infinite energy³⁶⁻³⁷. Many developments have been employed in recent years to solve this problem by including counterions from solutions to neutralize the net charge on the periodic reactive surface³⁸⁻⁴³.

3. Overview of the thesis

In electrochemistry, electrolytes are needed to transport charges between the cathode and the anode for conducting electricity. Since electrolytes are mobile ions in solution, heterogeneous catalysts in electrochemistry are typically in the solid phase. Solid catalysts can take many different forms, including perfect crystals, polycrystalline films or particles, and amorphous structures, with increasing complexity. Because chemical reactions take place at the interface between the heterogeneous catalysts and the reactants in the solution, the surface morphology directly determines the interfacial chemistry. However, because the complexities of the surface morphology differ in different categories of materials, the study of each type of catalysts require specific treatments. In this thesis, we employed quantum mechanical methods to study the three main categories of catalysts, as shown in Figure 1 below:

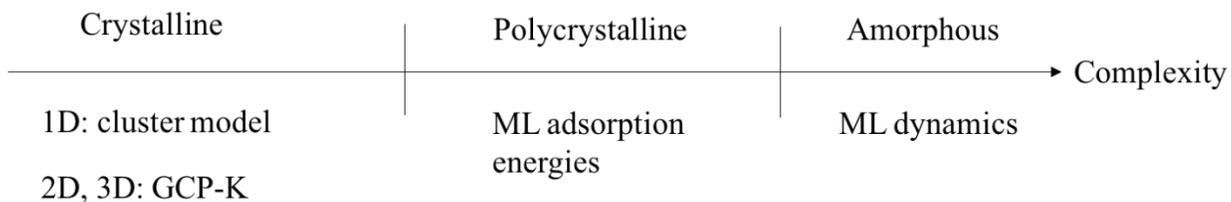


Figure 1. Systems and the corresponding methods studied in this thesis. GCP-K: grand canonical potential-kinetics; ML: machine learning.

First, quantum mechanical calculation is applied to study the simple low-index surface of crystalline materials. For extended systems, usually the periodic boundary condition is used, such that compounds in the forms of chains, slabs, or bulks are described by their dimensions as 1D, 2D, and 3D systems. The cluster model is first used on the 1-dimensional edge of crystalline MoS₂. In the calculation, the catalytic region of the linear edge is extracted to form a cluster by including a sufficient number of nearby atoms. Since the catalytic site is isolated from the extended system as a cluster, it can be treated as a homogeneous catalyst such that well-established molecular calculations can be used on this cluster model to study the catalysis. Using this method, we placed net charges on the catalyst as we did for the homogeneous catalysts. We found that the hydrogen evolution reaction takes place via the Heyrovsky mechanism involving a molybdenum hydride and a solvated hydronium water cluster. This work is summarized in Chapter 2 of the thesis.

Although the cluster model is useful for 1D systems, for higher dimensions, a cluster will contain a large number of atoms, making the method less effective. Thus, a different strategy is needed for more general and higher dimensional systems.

In our study of 2D and 3D catalytic systems, we re-formulated a macroscopic theory called the grand canonical potential kinetics (GCP-K) model to provide a foundation to relate quantum mechanical calculations to macroscopic analysis of reaction kinetics. By taking the Legendre transformation of the Gibbs free energy using the charge-voltage conjugate pairs, we obtained the grand canonical potential (GCP) that is quadratically dependent on the applied voltage. Then, by extending GCP for transition states during electrochemical reactions, we found that the reaction barriers could be calculated from either the constant charge or the constant applied voltage conditions. As the consequence of the new theory, the quadratic dependence on the applied voltage for species on the surface of an electrochemical heterogeneous catalyst is different from the conventional linear relationships given by the Nernst equation and Butler-Volmer kinetics for isolated molecules or reaction intermediates. Using this new methodology, we elucidated the hydrogen evolution reaction mechanism on the sulfur vacancy of MoS₂, as described in Chapter 3.

Although perfect crystalline surfaces are the simplest surfaces to study and to provide insights about the reactivity in certain chemical environments, under experimental conditions, more complex structures like polycrystalline and thin film structures are synthesized and used. Thus, to directly relate quantum mechanical calculations with actual experiments, theoretical calculations must be extended to consider various irregular morphologies on the surface.

To study the surface sites on polycrystalline systems, we combined the reactive force field (ReaxFF), quantum chemical calculations, and machine learning approaches. An important example of such systems is the copper nanoparticles on which the electrochemical CO₂ reduction is driven favorably towards C₂₊ products. The first difficulty in studying polycrystalline systems is to computationally generate a realistic particle for study. Here, the ReaxFF is specifically trained to match the dynamics of copper to simulate a realistic particle. With this nanoparticle, we can sample the surface sites for interactions with reaction intermediates using the cluster model. However, for a nanoparticle with 200,000 atoms and 20,000 surface sites, it is too computationally expensive to sample every site using quantum mechanics. Thus, an efficient machine learning algorithm is developed to predict the relevant properties of all the surface sites using only hundreds or thousands of random surface sites. Using this approach, we found that the favorable C₂₊ selectivity is due to the presence of a twin boundary site near the step surfaces involving the (100) and (111) surfaces, and we designed a new catalyst that will be more selective than the developed copper nanoparticles.

The above strategy for polycrystalline structure can easily be extended to amorphous structures. However, for the machine learning model to be self-consistent, in addition to be able to predict the adsorption energies, machine learning can also be used to model dynamical effects. For amorphous systems, accurate dynamics can be used to sample the complicated surface sites and their chemical environments. In this part of the thesis, the machine learning method is used to fit the potential energy surface (PES) of the system, such that forces are obtained by differentiating the total energy with respect to the atomic coordinates. We developed a new method to partition the total energy to each individual

atom for effective training of the neural network model. We showed that we could obtain errors as low as 1meV/atom in energy and 0.1eV/A in force using 10 times less data than the conventional approaches. To verify the methodology, we implemented and performed molecular dynamics simulation using the trained machine learning force field to calculate the heat conductivity of amorphous silicon, which requires a long molecular dynamics simulation that is not possible in typical quantum mechanical calculations. Such methodology can be easily extended to catalytically relevant systems.

In summary, we have used quantum mechanics and machine learning to study a wide range of heterogeneous catalytic systems, from simple crystalline structure, to polycrystalline structure, and to amorphous structures. In the development of these new approaches, we found that the surface electrochemistry is governed by more complex quadratic relations, rather than the simple Nernst equation and the Butler-Volmer type kinetics for homogeneous systems. In addition, machine learning methods have been shown to be promising to extend quantum mechanical results for more complex systems. We believe that the novel approaches used in the above systems will enable new insights to guide further developments of electrochemistry.

References

1. Pimentel, D.; Pimentel, M. H. *Food, Energy, and Society*, CRC Press, Taylor & Francis Group: Florida, 2008.
2. Rosa, A. E.; Machlis, G. E.; Keating, K. M. Energy and society. *Ann. Rev. Sociol.* **1988**. 14:149-72.

3. United States. Dept. of Energy. Office of Energy Analysis. *Annual Energy Outlook 2018*. Washington: Feb. 2018, Web.
4. BP. *Statistical Review of World Energy 2018*. Jun. 2018, Web.
5. World Energy Council. *World Energy, Issues Monitor 2018*. Dec. 2018, Web
6. Bithas K., Kalimeris P. (2016) A Brief History of Energy Use in Human Societies. In: *Revisiting the Energy-Development Link*. Springer Briefs in Economics. Springer, Cham: 2015.
7. Ourisson, G.; Albrecht, P.; Rohmer, M. The microbial origin of fossil fuels. *Scientific American*. **1984**. 2. pp 44-51.
8. Berner, R. A. The long-term carbon cycle, fossil fuels, and atmospheric composition. *Nature*. **2003**. 426, 323-326.
9. Zou, C.; Zhao, Q.; Zhang, G.; Xiong, B. Energy revolution: From a fossil energy era to a new energy era. *Natural Gas Industry B*. **2016**. 3(1), 1-11.
10. Norton, B. *Harnessing Solar Heat*. Springer Netherlands: The Netherlands, 2014.
11. Rodionova, M.V. et al. Biofuel production: challenges and opportunities. *International Journal of Hydrogen Energy*, **2017**. 42. 8450-8461.
12. Piotrowiak, P. *Solar energy conversion*, The Royal society of chemistry: Cambridge, UK. 2013.
13. Renewable Energy Policy Network for the 21st Century. *Renewables 2018, global status report*. Dec. 2018, Web.
14. Rahn, C. D.; Wang, C.-Y. *Battery systems engineering*. John Wiley & Sons, Ltd. U.K. 2013.

15. O'Hayre, R.; Cha, S.-W.; Colella, W. G.; Prinz, F. B. *Fuel cell fundamentals, third edition*. John Wiley & Sons. New Jersey 2016.
16. Vangari, M.; Pryor, T.; Jiang, L. Supercapacitors: review of materials and fabrication methods. *Journal of Energy Engineering*, **2013**, 139(2). pp 72.
17. Detz, R. J.; Reek, J. N.; van der Zwaan, B. C. C. The future of solar fuels: when could they become competitive? *Energy Environ. Sci.*, **2018**, 11, 1653.
18. Crabtree, G. The hydrogen economy. *Physics Today*, **2004**, 57(12), 39.
19. Spurgeon, J. M.; Kumar, B. A comparative technoeconomic analysis of pathways for commercial electrochemical CO₂ reduction to liquid products. *Energy Environ. Sci.*, **2018**, 11, 1536.
20. Morrison, S. R. *Electrochemistry at Semiconductor and Oxidized Metal Electrodes*, Springer U.S., Plenum Press, New York: 1980.
21. Bard, A. J., Faulkner, L. R. *Electrochemical Methods*, John Wiley & Sons, New York: 2001.
22. Sato, N. *Electrochemistry at Metal and Semiconductor Electrodes.*, Elsevier Science B. V., Amsterdam, The Netherlands: 1998.
23. Trogadas, P.; Ramani, V.; Strasser, P.; Fuller, T. F.; Coppins, M.-O. Hierarchically Structured Nanomaterials for Electrochemical Energy Conversion. *Angew. Chem. Int. Ed.* **2016**, 55, 122-148
24. Kolasinski, K. W. *Surface science*. John Wiley & Sons. Ltd. United Kingdom: 2012.
25. Szabo, A.; Ostlund, N. S. *Modern quantum chemistry*. McGraw-Hill, Inc., New York 1989.

26. Jensen, F. *Introduction to computational chemistry*. John Wiley & Sons, Ltd. England 2007.
27. Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular electronic-structure theory*. John Wiley & Son Ltd: England, 2000.
28. Popelier, P. *Solving the Schrödinger equation*. Imperial College Press: London, 2011.
29. Hohenberg, P; Kohn, W. Inhomogeneous electron gas. *Physical Review*, **1964**. 136(3B): B864-B871.
30. Engel, E.; Dreizler, R. M. *Density Functional Theory*. Springer-Verlag Berlin Heidelberg: 2011.
31. Kohn, W.; Sham, L. J. Self-consistent equations including exchange and correlation effects. *Physical Review*, **1965**. 140(4A): A1133-A1138.
32. Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1997**. 78, 1396.
33. Becke, A. D. Density-functional thermochemistry. III. The role of exact exchange. *J. Chem. Phys.* **1993**. 98, 5648.
34. Hill, T. L. *An introduction to statistical thermodynamics*. Courier Corporation: United States, 1986.
35. Skyner, R. E.; McDonagh, J. L.; Croom, C. R.; van Mourik, T.; Mitchell, J. B. O. A review of methods for the calculation of solution free energies and the modelling of systems in solution. *Phys. Chem. Chem. Phys.* **2015**, 17, 6174-6191.
36. Griffiths, D. J. *Introduction to Electrodynamics*. Prentice Hall: New Jersey: 1981.

37. Jackson, J. D. Classical electrodynamics, third edition. John Wiley & Sons: Singapore, 1999.
38. Skúlason, E., Karlberg, G. S., Rossmeisl, J., Bligaard, T., Greeley, J., Jónsson, H., Nørskov, J. K. Density functional theory calculations for the hydrogen evolution reaction in an electrochemical double layer on the Pt(111) electrode. *Physical Chemistry Chemical Physics*, **2007**, 9(25), 3241–3250.
39. Taylor, C. D., Neurock, M. Theoretical insights into the structure and reactivity of the aqueous/metal interface. *Current Opinion in Solid State and Materials Science*, **2005**, 9(1–2), 49–65.
40. Taylor, C., Wasileski, S., Filhol, J.-S., Neurock, M. First principles reaction modeling of the electrochemical interface: Consideration and calculation of a tunable surface potential from atomic and electronic structure. *Physical Review B*, **2006**, 73(16), 165402.
41. Jinnouchi, R., Anderson, A. B. Electronic structure calculations of liquid-solid interfaces: Combination of density functional theory and modified Poisson-Boltzmann theory. *Physical Review B*, **2008**, 77(24), 245417.
42. Sha, Y.; Yu, T. H.; Merinov, B. V.; Goddard, W. A. Prediction of the Dependence of the Fuel Cell Oxygen Reduction Reactions on Operating Voltage from DFT Calculations. *The Journal of Physical Chemistry C*, **2012**, 116(10), 6166–6173.
43. Gunceler, D., Letchworth-Weaver, K., Sundararaman, R., Schwarz, K. a, Arias, T. A. The importance of nonlinear fluid response in joint density-functional theory studies of battery systems. *Modelling and Simulation in Materials Science and Engineering*, **2013**, 21(7), 074005.

THE REACTION MECHANISM WITH FREE ENERGY BARRIERS FOR ELECTROCHEMICAL DIHYDROGEN EVOLUTION ON MoS₂

Abstract

We report Density Functional Theory (M06L) calculations including Poisson-Boltzmann solvation to determine the reaction pathways and barriers for the hydrogen evolution reaction (HER) on MoS₂, using both a periodic two-dimensional slab and a Mo₁₀S₂₁ cluster model. We find that the HER mechanism involves protonation of the electron rich molybdenum hydride site (Volmer-Heyrovsky mechanism), leading to a calculated free energy barrier of 17.9 kcal/mol, in good agreement with the barrier of 19.9 kcal/mol estimated from the experimental turnover frequency. Hydronium protonation of the hydride on the Mo site is 21.3 kcal/mol more favorable than protonation of the hydrogen on the S site because the electrons localized on the Mo-H bond are readily transferred to form dihydrogen with hydronium. We predict the Volmer-Tafel mechanism in which hydrogen atoms bound to molybdenum and sulfur sites recombine to form H₂ has a barrier of 22.6 kcal/mol. Starting with hydrogen atoms on adjacent sulfur atoms, the Volmer-Tafel mechanism goes instead through the M-H + S-H pathway. In discussions of metal chalcogenide HER catalysis, the S-H bond energy has been proposed as the critical parameter. However, we find that the sulfur-hydrogen species is not an important intermediate since the free energy of this species does not play a direct role in determining the effective activation barrier. Instead we suggest that the kinetic barrier should be used as a descriptor for reactivity, rather than the equilibrium thermodynamics. This is supported by the agreement between the calculated barrier and the experimental turnover frequency. These results suggest that to design a more reactive catalyst from edge exposed MoS₂, one should focus on lowering the reaction barrier between the metal hydride and a proton from the hydronium in solution.

1. Introduction

A major challenge in renewable energy technology is water splitting; that is, using solar radiation to photoelectrochemically convert water molecules into H₂ and O₂. Here, both the hydrogen evolution reaction (HER) and the oxygen evolution reaction (OER) present challenges for the catalysts. The detailed reaction mechanisms have not yet been established for either one. Here, we consider the easier case of HER.

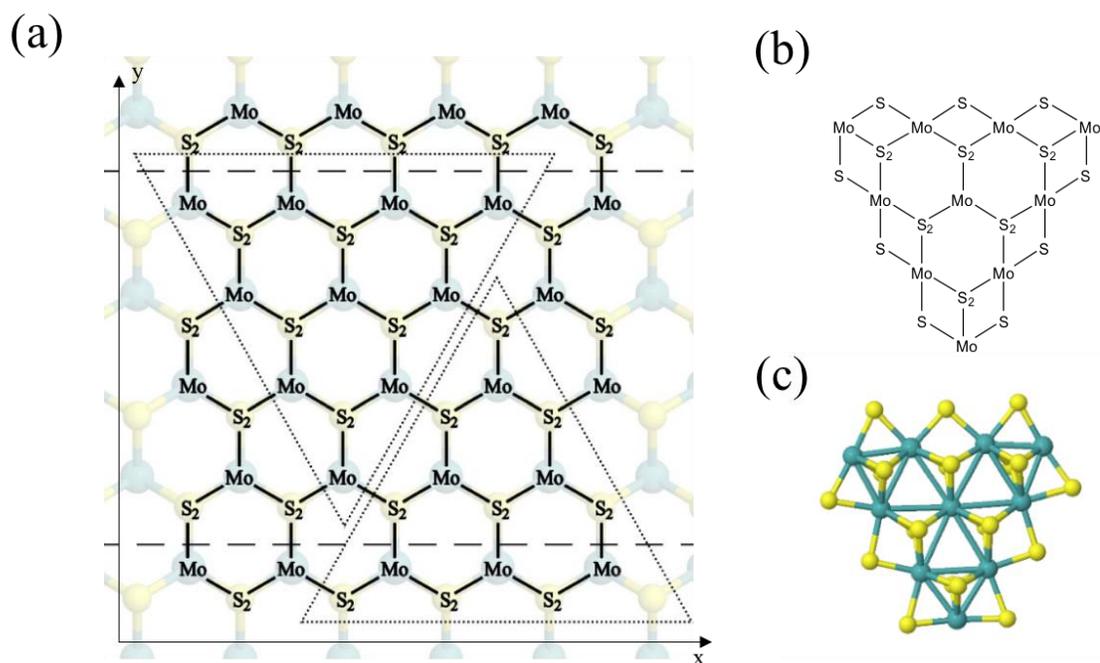
Platinum is the most efficient inorganic HER catalyst; however, the cost of platinum has motivated an extensive search for earth abundant HER catalysts¹. Hinnemann et al.² reported DFT calculations suggesting that molybdenum disulfide (MoS₂) can catalyze HER because of its nearly thermoneutral hydrogen adsorption energy. Indeed, Jaramillo et al.³ synthesized MoS₂ on Au(111) and showed that the activity for HER correlates linearly with the total length of the exposed edges of crystalline MoS₂. Since then, numerous experimental synthetic techniques have attempted to synthesize MoS₂ catalysts that carry out HER more efficiently⁴. Molecular analogues were also studied⁵.

Although the hydrogen adsorption energy has been a useful descriptor for screening materials to identify candidates for HER, it is reaction barriers that determine the rates. Thus, to design the most efficient HER catalysts, we must determine the reaction barriers for the various reaction sequences that can convert protons and electrons to H₂. Here, the pathway with the lowest rate-determining step (RDS) is expected to dominate the reaction rate.

In this study, we perform Density Functional Theory (DFT) quantum mechanics calculations to determine the reaction pathway for HER on the Mo-edge (10 $\bar{1}$ 0) of MoS₂. To enable the use of the most accurate DFT for reaction barriers while describing solvation effects at the Poisson Boltzmann level, we describe the Mo-edge of MoS₂ using a Mo₁₀S₂₁ cluster model. This allows us to consider the introduction of protons and electrons separately and report free energies as a function of electrochemical potential and pH.

2. Computational Methods

2.1 Finite Cluster. For the finite clusters, we used the M06L⁶ flavor of DFT, which has been found to give reliable energy barriers for reaction mechanisms of organometallic catalysts⁷. In these calculations, we use the small core angular momentum projected effective core potential⁸ from Hay and Wadt⁹ to replace the inner 28 core electrons, so that 14 explicit electrons are described on each Mo (denoted LACVP**). All electrons are included for each S and H using the 6-31G** basis set¹⁰ for structure optimization, and the 6-311G**++ basis set^{11,12} (hydrogen and oxygen) and the 6-311G-3df basis set^{12,13} (sulfur) for the final electronic energy. The solvation energy was calculated by solving the Poisson-Boltzmann equation (PBF)¹⁴ at the optimized geometry (using a solvent radius of 1.4Å and a dielectric constant of 80.37). The vibrational frequencies used in obtaining the zero point energy, entropy, and enthalpy were calculated using the CPHF technique at the optimized geometry. All cluster calculations were performed with JaguarTM 15.



(d)

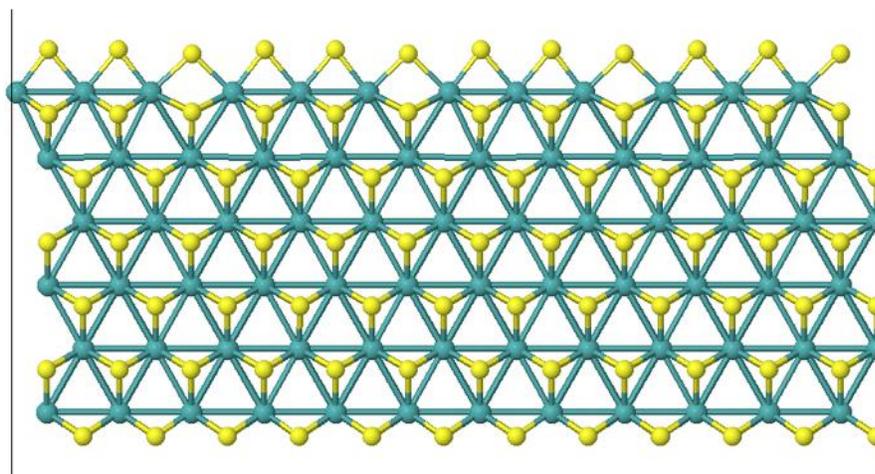


Figure 1. (a) Top view of the 2D MoS₂ sheet. The two horizontal dashed lines indicate terminations along the (10 $\bar{1}$ 0) Mo-edge and ($\bar{1}$ 010) S-edge. The two triangles represent the terminations for Mo-edge and S-edge clusters; (b) Schematic representation of the 50% S coverage Mo-edge cluster; (c) Optimized triangular Mo-edge cluster with stoichiometry of Mo₁₀S₂₁. (d) Optimized structure of a 12Mo x 6Mo slab, where the Mo-edge (top edge) clusters into groups of 3 Mo. The same structure is obtained for the 3Mo and 6Mo wide slab.

The free energy of an H₂ molecule at 1atm and 298K was calculated as above. The free energy of H⁺ at 1M in water ($G = -270.3$ kcal/mol) was taken to be its gas-phase value ($G(\text{H}^+, 1\text{atm}) = H - TS = 2.5 k_B T - T * 26.04 = -6.3$ kcal/mol) plus the empirical hydration energy ($G(\text{H}^+, 1\text{atm} \rightarrow 1\text{M}) = -264.0$ kcal/mol)¹⁶. Using the definition of the standard hydrogen electrode (SHE) condition where e⁻ and H⁺ (pH=0) are in equilibrium with 1atm H₂, the free energy of an electron at SHE can be determined as the difference between the free energies of $\frac{1}{2}\text{H}_2$ and H⁺. The chemical potential of electrons and protons away from the SHE condition are then calculated as

$$\mu_e(E) = \mu_e(\text{SHE}) - 23.06 \times E, \text{ and}$$

$$\mu_H(\text{pH}) = \mu_H(\text{pH}=0) - 1.36 \times \text{pH}, \text{ both in kcal/mol.}$$

With these reference values, the free energies of the cluster with various numbers of electrons and protons can be compared.

2.2 Periodic Slab. For the periodic slab calculations we consider a single MoS₂ slab terminated on the (10 $\bar{1}$ 0)(Mo-edge) and ($\bar{1}$ 010) (S-edge) boundaries with three Mo per unit cell as shown in Figure 1. Here we used the PBE¹⁷ flavor of DFT with a plane wave basis set extending to 300eV for geometry optimization and 500eV for the final electronic energies. We used the projected augmented wave (PAW)¹⁸ method for pseudopotentials to represent the effects of the inner 36 core electrons of Mo and the 10 electrons of the S, so that only the outmost 6 electrons of both Mo and S are described explicitly. Slab calculations are performed in VASP¹⁹.

Since the slab model consists of only a single S-Mo-S tri-layer with both the Mo-edge and the S-edge exposed, we used the Monkhorst-Pack $4 \times 1 \times 1$ k-point mesh, where 4 k-points are used in the periodic direction. The electronic energy was minimized within 0.1 mH, and the geometry was optimized to 0.1 mH/Å.

For these periodic calculations, the reference H₂ molecule is calculated in a $15\text{Å} \times 15\text{Å} \times 15\text{Å}$ cubic vacuum space with only the Γ point.

3. Model Systems

3.1 Periodic Slab. The dominant phase for bulk crystalline MoS₂ is the 2H phase with AB stacking of the graphene-like hexagonal 2D sheets²⁰. Upon synthesis, the exposed surfaces are generally the (001) basal plane of the S-Mo-S tri-layer, the Mo-edge (10 $\bar{1}$ 0) and S-edge ($\bar{1}$ 010). It was shown experimentally that the activity of the catalyst correlates directly with the total length of the exposed edges³. A recent experiment confirms this observation by comparing the activities between the edges and the basal plane²¹. In addition, we expect the chemistry of the edges of bulk MoS₂ to be similar to that of a single MoS₂ layer. It has been shown that additional layers decrease the current density due to electron hopping across the layers, such that the top layers are not as active as the bottom layers²². The exchange current

density ($0.6 \times 10^{-7} \text{ A/cm}^2$)²² reported for triangular pyramidal MoS₂ platelets is in the same order of magnitude as monolayer MoS₂ ($1.3 \times 10^{-7} \text{ A/cm}^2$)³. The edges from the top layers contribute little to the activity, thus the monolayer MoS₂ is sufficient to understand the chemistry of HER on MoS₂ edges.

STM studies of MoS₂ crystals show three stable edge configurations: 100% S coverage of the Mo-edge, 50% S coverage of the Mo-edge, and the S-edge²³. Under the sulfiding conditions appropriate for hydrodesulfurization (HDS) processes, the Mo-edge has been established experimentally as the dominant form; however, under HER conditions with no external sulfur = 50% S coverage of the Mo-edge is the most stable form²⁴⁻²⁶. This has also been observed in a recent HR-STEM study²⁷. This structure was assumed in the previous DFT studies of HER on MoS₂². Thus, to study HER electrocatalysis, we chose to examine the 50% coverage case for the Mo-edge.

Experimental studies have observed finite MoS₂ clusters ranging from 1 nm to bulk. However, for systems of such large sizes, it is convenient to use periodic boundary conditions in the x direction, exposing the Mo-edge in the +y direction and the S-edge in the -y direction (Figure 1). We carried out DFT calculations for such periodic slabs and found that models that are three or six unit cells long in the periodic direction relax to structures 0.1 eV/Mo lower than models that are either two or four unit cells long. This has also been observed in other studies of the 50% S coverage Mo-edge calculations but not yet explained^{25,28}.

This 3-Mo periodicity can be understood in terms of simple electron counting arguments. In the bulk, the oxidation state of each Mo atom is +4 with bonds to 6 sulfur atoms, each of which bonds to 3 Mo. Thus we can consider there to be $2/3$ electrons from each Mo in each bond. In the valence bond description, this is described in terms of resonating structures. However, the surface stabilizes Mo-S valence bond structures with local 2-electron bonds. Thus, for the Mo terminated surface each edge sulfur contributes 1 electron to each Mo-S bond. Therefore each edge Mo contributes $2 \times 1 + 4 \times 2/3 = 14/3$ electrons to its six S neighbors. This implies that each triad of 3 Mo atoms on the edge contributes a total of 14 electrons for bonding. This would correspond to a d² configuration on one Mo and a d¹ configuration on

the other two. Alternatively, one can say that the full d2 band of the bulk system leads to a surface d2 band that is only 2/3 full. This leads to a Peierls distortion that has a periodicity of 3. Therefore, periodic calculations for slabs should always have a multiple of 3 edge Mo slab because it leads to an integer number of bonding electrons on the edge. In this reconstruction, three consecutive Mo atoms draw together (Mo-Mo distances of 2.96 Å), leaving a larger separation (3.59 Å) between the triads. Indeed, the calculations on an edge with 6 and 12 Mo periodicity in the x- direction show the same triad reconstruction, as shown in Figure 1d.

3.2 Cluster Model. Using molecular clusters to model a periodic system for determining reaction mechanisms allows more flexibility in the accuracy of the methods (allowing us to use M06L, which is more accurate for reaction barriers and bond energies than PBE⁷, the most common method for periodic calculations). It is straightforward to use clusters with net charges (difficult in PBC), and we can use the PBF (Poisson Boltzmann) continuum solvation method that has been well-validated for aqueous solvation energies²⁹. Indeed, experiments can be carried out on supported MoS₂ to validate predictions.

Figure 1a shows how we extract a triangular cluster from the periodic array to expose only the Mo edges. Figure 1b shows the schematic representation of the finite cluster and Figure 1c shows the optimized structure.

3.3 Validation of the Cluster Model. To validate that the cluster model has the same chemical properties as the periodic Mo-edge, we calculated the binding energy of a hydrogen atom to both the cluster and the periodic slab, both under vacuum conditions. In both cases, we reference the free H atom energy to that of 1/2 H₂ molecule. The results in Figure 2 show that the bond energies calculated using the cluster and slab models differ by 0.03, 0.06, 0.01, and 0.03 eV for the four stoichiometries.

Reduction

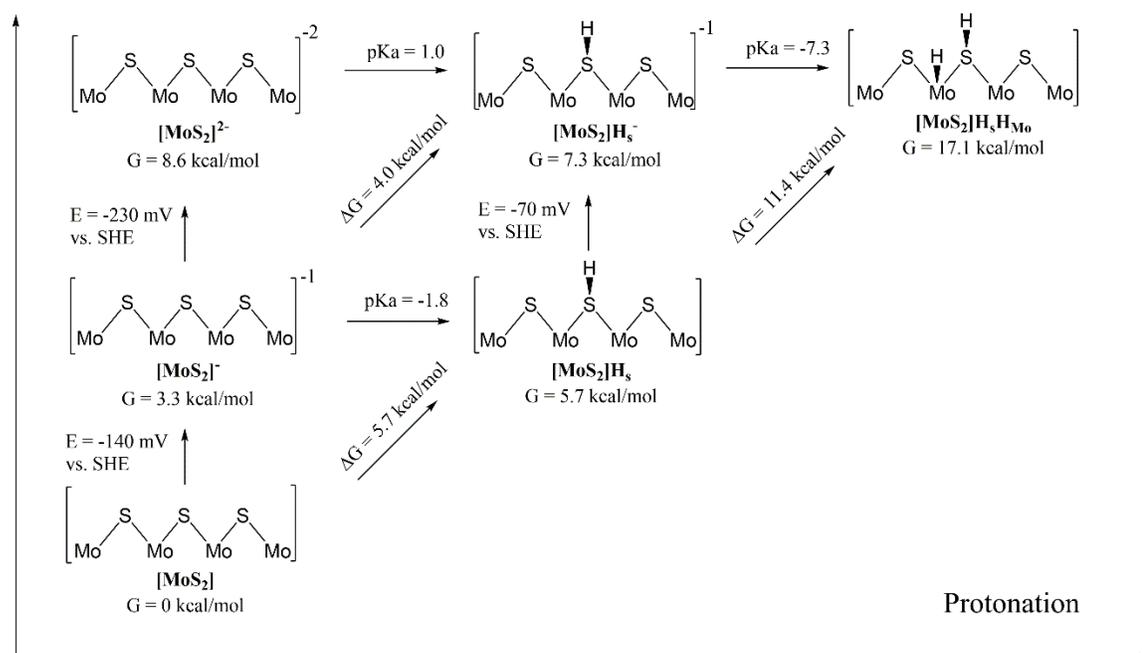


Figure 3. Thermodynamics (free energies at 298K, pH=0, and E=0V vs. SHE) for the most stable configurations as a function of the number of protons and electrons added. Here we start with the Mo-edge cluster at equilibrium at the SHE potential. Each structure to the right has one additional proton. Each structure along the ordinate has one more electron, representing a reduction step.

In order to evolve an H_2 molecule, protons and electrons must be added to the cluster. Here, it is useful to examine first the most stable structures with each number of extra electrons and each number of extra protons to understand the free energy differences between intermediate states, and ultimately find the lowest-barrier pathway. These free energies are shown in Figure 3.

1. At SHE conditions with $E=0\text{V}$ and $\text{pH}=0$, the most stable state is $[\text{MoS}_2]$, the bare neutral Mo-edge.
2. The first reduction potential to obtain $[\text{MoS}_2]^-$, is only -140mV , leading to a negatively charged cluster solvated in water.

3. Protonating $[\text{MoS}_2]$ leads to $[\text{MoS}_2]\text{H}_\text{s}^+$ (subscript s indicates the hydrogen atom is bound to a S atom) with an energy cost of 7.2 kcal/mol.
4. Instead of Step 3, we can add a proton to the edge S having an extra electron, $[\text{MoS}_2]^-$, to form $[\text{MoS}_2]\text{H}_\text{s}$, which costs $\Delta G=2.4$ kcal/mol, corresponding to a pK_A of -1.8. Comparing to $[\text{MoS}_2]$ the cost of adding an H atom (a proton and an electron simultaneously) is $\Delta G=5.7$ kcal/mol as in Figure 2. If instead we put the H on the Mo to obtain $[\text{MoS}_2]\text{H}_\text{Mo}$, the energy is 25.7 kcal/mol relative to $[\text{MoS}_2]$, corresponding to an excitation energy of $\Delta G=20.0$ kcal/mol.
5. Adding now an electron to $[\text{MoS}_2]\text{H}_\text{s}$ to form $[\text{MoS}_2]\text{H}_\text{s}^-$ occurs at a potential of -0.07 V. Moving the H from the S to the Mo, to form $[\text{MoS}_2]\text{H}_\text{Mo}^-$, costs $\Delta G=5.9$ kcal/mol.
6. Starting with $[\text{MoS}_2]\text{H}_\text{s}^-$ from Step 5 we can add a proton to obtain $[\text{MoS}_2]\text{H}_\text{s}\text{H}_\text{Mo}$ at a cost of 9.7 kcal/mol ($\text{pK}_\text{A} = -7.3$).
7. Thus, the $[\text{MoS}_2] \rightarrow [\text{MoS}_2]^- \rightarrow [\text{MoS}_2]\text{H}_\text{s} \rightarrow [\text{MoS}_2]\text{H}_\text{Mo}^- \rightarrow [\text{MoS}_2]\text{H}_\text{s}\text{H}_\text{Mo}$ pathway may proceed with no highly endergonic steps.

We can also consider the case with more hydrogen on the cluster, which would be required by a Volmer-Tafel mechanism. However, as shown in Figure 3, an extra hydrogen is thermodynamically much more difficult, with a cost of about 11 kcal/mol ($[\text{MoS}_2]\text{H}_\text{s} \rightarrow [\text{MoS}_2]\text{H}_\text{s}\text{H}_\text{Mo}$).

Based on this thermochemistry, the highest-free energy intermediates for either a Heyrovsky ($[\text{MoS}_2]\text{H}^-$) or Tafel ($[\text{MoS}_2]\text{H}_\text{s}\text{H}_\text{Mo}$) mechanism are the last intermediates in the cycles. Therefore, we searched for rate-limiting transition states by considering the H-H bond forming steps, then confirmed that the barriers connecting the preceding lower-energy intermediates are indeed lower.

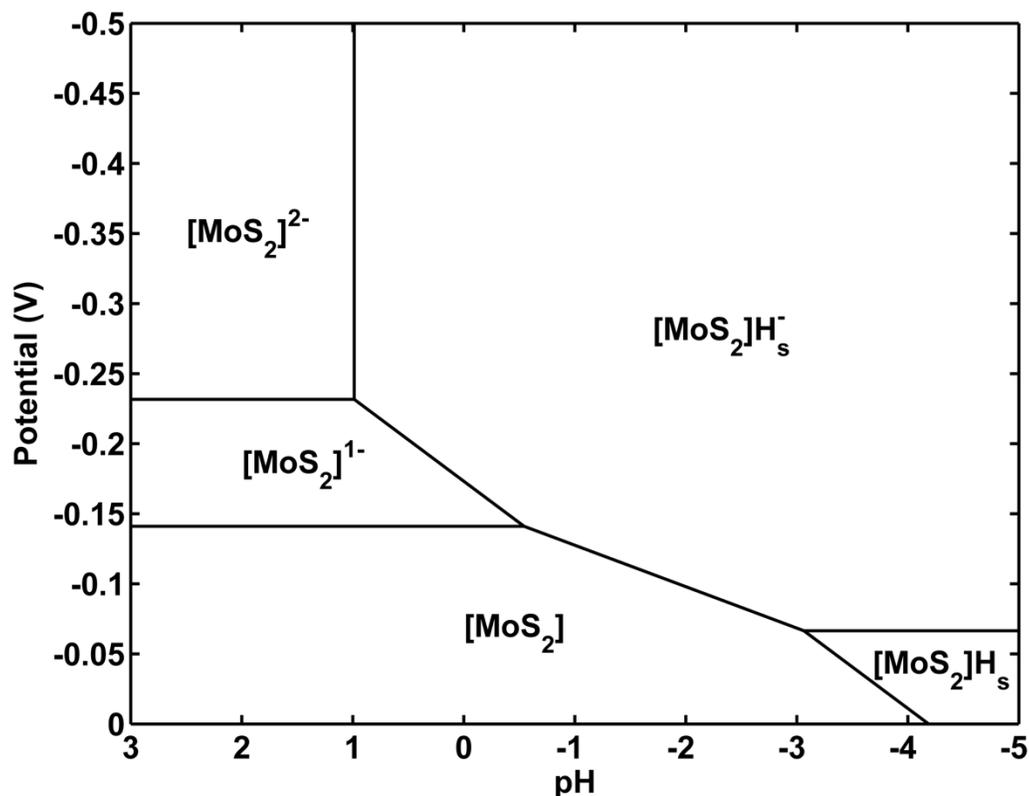


Figure 4. Calculated Pourbaix diagram for the surface states of the Mo-edge cluster.

4.2 Pourbaix Diagram. The intermediate species considered in Figure 3 lead to the Pourbaix diagram in Figure 4, showing the dominant phases as a function of pH and external potential. It is sufficient to use only the species in Figure 3, since only the most stable structure for a given stoichiometry appears on the Pourbaix diagram.

In the range of pH and potentials that are relevant to HER, 5 states are present. Starting from SHE at pH=0 and E=0V, and applying increasingly negative potentials, the cluster is reduced first to $[\text{MoS}_2]^-$ to become negatively charged. Further reducing the potential leads to the $[\text{MoS}_2]\text{H}_s^-$ structure (at pH=0) rather than the -2 charged state. At very negative pH and potential, the $[\text{MoS}_2]\text{H}_s$ structure is the most favorable, since the potential is not sufficiently strong to further reduce the structure.

4.3 Transition States Analysis. Two types of transition states are considered in this study: the Volmer-Tafel mechanism in which two adsorbed hydrogens react to form H_2 , and the Volmer-Heyrovsky mechanism with one adsorbed hydrogen reacting with a solvated proton. The calculated free energy barriers are shown in Figure 5, at a potential of 0V and pH = 0.

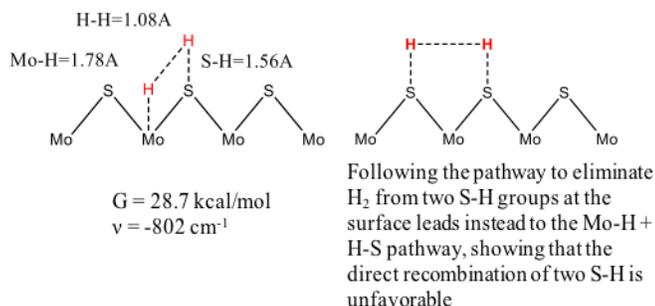
4.3.1 Volmer-Tafel Mechanism. In the Volmer-Tafel mechanism, two adsorbed hydrogens next to each other react to form a dihydrogen molecule. Since there are two types of atoms, S and Mo, on the edge, we considered reactions of the type S-H + H-S and Mo-H + H-S. However, constraining the H atoms on two adjacent sulfurs to move toward a possible transition state, the hydrogen on one sulfur instead first moves onto the nearby Mo atom, forming a Mo-H species. Thus, a transition state of the type S-H + H-S is electronically unreasonable. Thus, the Volmer-Tafel reaction on the Mo-edge takes place between hydrogens on adjacent Mo and S atoms. Beginning from a relaxed geometry in which the H-H distance was constrained to 1.0 angstrom, eigenvector following was used to locate a transition state with a single imaginary frequency. This transition state geometry was found to have bond distances of

- 1.78 Å for Mo-H (compared to the equilibrium Mo-H bond of 1.67Å);
- 1.08 Å for H-H (compared to a final H-H bond of 0.74Å), and
- 1.56 Å for S-H (compared to a normal S-H bond of 1.35Å).

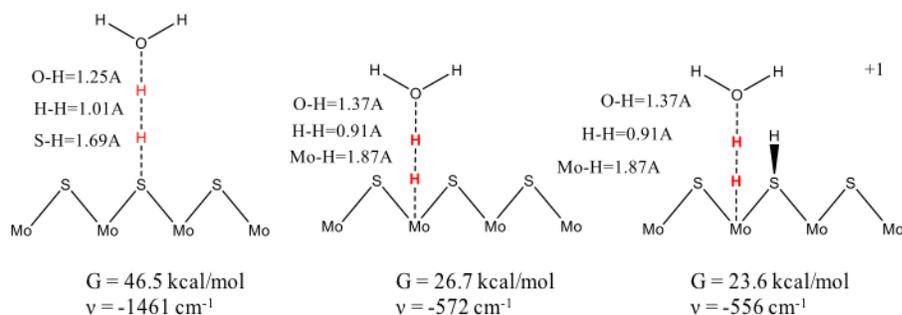
The free energy at the transition state is 28.7 kcal/mol relative to the $[MoS_2]$ ground state. It is 11.6 kcal/mol above the preceding intermediate $[MoS_2]H_S H_{Mo}$.

4.3.2 Volmer-Heyrovsky Mechanism. The Volmer-Heyrovsky mechanism is more complicated because it is necessary to solvate the H_3O^+ source of the proton along the reaction pathway. We find that obtaining accurate results requires the use of a cluster of 4 waters, one of which is protonated at the beginning but all of which are neutral at the end. However, in the reaction between an adsorbed hydrogen atom and the hydronium bound proton, the water cluster must rearrange to expose the proton for reaction.

(a) Barriers for the Volmer-Tafel mechanisms:



(b) Barriers for the Volmer-Heyrovsky mechanisms:



(c) Hydronium water clusters used in studying the Heyrovsky reactions (the full system with 4 explicit H_2O was solvated using PBF):

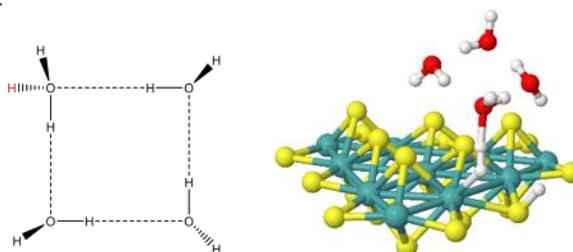


Figure 5. Schematics of the transition state structures considered for H_2 formation. Bond lengths and imaginary frequencies confirm transition state searches did not relax to products. The hydronium water cluster used for the Heyrovsky step is also shown. The red **H** indicates the reacting hydrogen atoms.

First, we examined the hydronium reacting with an adsorbed hydrogen atom bound to the edge sulfur atom. Again, transition state structures were located by following imaginary modes from constrained initial guesses. The transition state found for S-H + H-OH₂ has a free energy of 46.5 kcal/mol, which is too high for hydrogen evolution on the Mo-edge.

Using the same strategy, we found that the transition state free energies for reacting with the Mo hydrides [MoS₂]H_{Mo}⁻ and [MoS₂]H_{Mo}H_S are 26.7 and 23.6 kcal/mol. The barriers for this process are much lower than the S-H case, making them the most favorable transition structures to form H₂. The free energy profile for these reactions shows that protonation of the edge S atom promotes hydride transfer, resulting in a lower barrier for hydride transfer from [MoS₂]H_{Mo}H_S (G=17.1kcal/mol) than from [MoS₂]H_{Mo}⁻ (G=13.2kcal/mol). The bond distances for the [MoS₂]H_{Mo}H_S case are 1.93 Å for Mo-H, 0.85 Å for H-H, and 1.51 Å for H-OH₂, which correspond to a late transition state.

4.3.3 Overall Reaction. For the Mo-H + H-OH₂ transition state to be rate-limiting, we must confirm that the barriers in the previous steps are smaller. Choosing the [MoS₂] → [MoS₂]⁻ → [MoS₂]H_S → [MoS₂]H_{Mo}⁻ → [MoS₂]H₂ path described in the thermodynamic section, we calculated transition states for the protonation and migration steps that are of smaller energy than the Heyrovsky barrier, as shown in Figure 6.

This mechanism gives insights on why MoS₂ is a good catalyst. The direct protonation of the Mo atom leads to a calculated high barrier of 28.1 kcal/mol, but in the presence of the S atom, the hydrogen adsorbs first on the chalcogenide, then migrates to the Mo atom (barrier = 20.5 kcal/mol), and finally reacts with a proton from solution to form H₂. These multiple steps lower the barrier for the whole process.

Toulhoat et al. computationally examined the dissociation of H₂ on MoS₂ edges in the context of hydrodesulfurization (i.e., water free)³⁰. In the absence of a protic solvent, our results are consistent that, among reaction mechanisms for the cleavage/formation of the H-H bond, the Tafel reaction of the type Mo-H + H-S provides the lowest energy pathway.

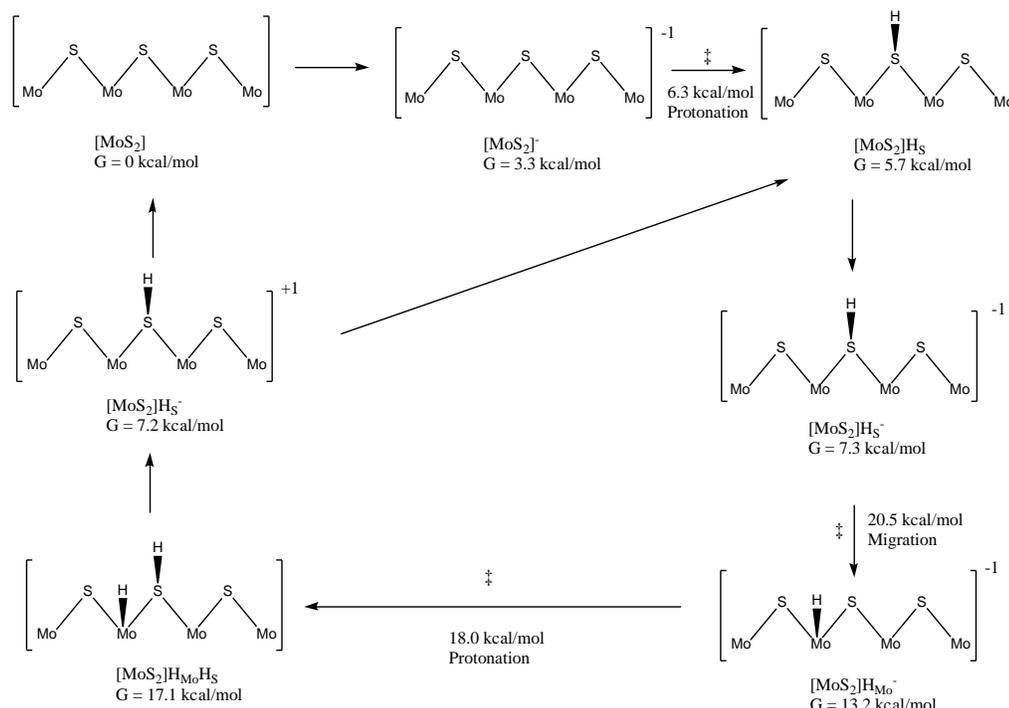


Figure 6. Overall reaction mechanism for HER on the Mo-edge cluster. The edge is first reduced by one electron, followed by a protonation to create a sulfur-bound hydrogen with a small barrier of 6.3 kcal/mol. The structure is further reduced by one more electron, followed by the migration of hydrogen onto the Mo atom with a relatively high barrier of 20.5 kcal/mol. The edge is protonated again by hydronium with a barrier of 18.0 kcal/mol. Finally, the metal hydride reacts with a proton from solution, forming dihydrogen with a barrier of 23.6 kcal/mol. All the free energies are relative to the ground state with no adsorbed hydrogens.

Finally, our result can be compared with estimates from Jaramillo's experimental study³. At a potential of -150mV and pH = 0.24, he estimated a turnover frequency (TOF) of $1.64 \times 10^{-2} \text{ s}^{-1}$ per edge molybdenum atom for hydrogen evolution on the Mo edge clusters. Using transition state theory,

$$\text{Rate} = (k_B T/h) \times \exp(-\Delta G^\ddagger/RT)$$

This TOF corresponds to a barrier of 19.9 kcal/mol. After adjusting our theoretical calculations for the chemical potentials due to the applied potential and pH, the ground state shifts to the negatively charged unprotonated structure, as shown in Figure 4. From the new resting state, we calculate an adjusted barrier of 17.9 kcal/mol, very close to the 19.9 kcal/mol from the experimental estimate.

The Tafel slope b can also be estimated from the theoretical calculation, assuming electron transfer from the support to the catalyst does not limit the rate. We expect that $b=2.3RT/nF\approx 60\text{mV}/n$, where n is the difference in the number of electrons between the ground state and the transition state. Under reaction conditions, $n=1$ since the ground state has shifted to the negatively charged structure, as shown in the Pourbaix diagram in Figure 4. In this case, the Tafel slope is 60 mV/dec, in agreement with the value of 55~60mV/dec from the experimental Tafel plot³. Different polymorphs of MoS₂ have been shown to have different Tafel slopes, likely due to the effects of conductivity, which are minimized in the case of single MoS₂ layers on Au(111). For example, Chhowalla et al. measured a lower Tafel slope of 40mV/dec using 1T-MoS₂, but proposed a different active site in the basal planes was responsible for HER³¹. Cao et al. also reported that the transition from crystalline MoS₂ to amorphous MoS₃ resulted in a change of Tafel slope from c.a. 90mV/dec to 40 mV/dec³².

5. Conclusion

Modeling the Mo-edge of a single MoS₂ sheet with a Mo₁₀S₂₁ cluster model, we found that the HER mechanism takes place through the Volmer-Heyrovsky mechanism involving an electron rich molybdenum hydride and a hydronium cation. This leads to an estimated barrier of 17.9 kcal/mol in good agreement with the experiment, in which the barrier 19.9 kcal/mol is estimated from the turnover frequency (TOF).

We find hydronium protonation of the hydride on the Mo site is 21.3 kcal/mol more favorable than protonation of the hydrogen on the sulfur because the electrons localized on the Mo-H bond are readily transferred to form dihydrogen with hydronium.

The S-H + H-S Volmer-Tafel mechanism and the S-H + H-OH₂ Volmer-Heyrovsky mechanisms attempt to combine highly acidic hydrogen (i.e., two protons) to form H₂, leading to a higher barrier for transfer of electrons from the bulk. The Volmer-Tafel mechanism between molybdenum hydride and sulfur hydrogen has a relatively low barrier of 28.7 kcal/mol, for the same reason that electrons are easily transferred from hydride to form the hydrogen-hydrogen bond. Indeed, it is possible that tuning the pK_A of the edge chalcogenide could lower the barrier for this mechanism.

Based on the volcano plot concept³³, it is widely believed that the binding energy of hydrogen on sulfur is the most important factor towards HER on crystalline MoS₂. However, we find that the sulfur-hydrogen species is not the critical intermediate since the free energy of this species does not play a role in determining the effective activation barrier.

In fact, we find that the predicted activation barrier for reaction between molybdenum hydride and hydronium is in agreement with the experimental rate of per-site TOF. This indicates that kinetic parameters should be used as a descriptor for reactivity, rather than equilibrium thermodynamics.

In conclusion, to design a more reactive catalyst from exposed MoS₂ edges, one should focus on lowering the reaction barrier between the metal hydride and a positively charged proton, either from the sulfur hydrogen species on the edge or from the hydronium in solution.

References

1. Laursen, A. B.; Kegnæs S.; Dahl, S.; Chorkendorff, I. *Energy Environ. Sci.* **2012**, *5*, 5577-5591. Merki, D.; Hu, X. *Energy Environ. Sci.* **2011**, *4*, 3878-3888. Yang, J.; Shin, H. S. *J. Mater. Chem. A.* **2014**, *2*, 5979-5985
2. Hinnemann, B.; Moses, P. G., Bonde, J.; Jørgensen, K. P.; Nielsen, J.H.; Horch, S., Chorkendorff, I.; Nørskov, J. K. *J. Am. Chem. Soc.* **2005**, *127*(15), 5308-5309.

3. Jaramillo, T. F.; Jørgensen, K. P.; Bonde, J. Nielsen, J. H.; Horch, S.; Chorkendorff, I. *Science*, **2007**, 317(5834), 100-102.
4. Kibsgaard, J.; Chen, Z.; Reinecke, N.; Jaramillo, T. F. *Nature Materials* **2012**, 11, 963-969. Li, Y.; Wang, H.; Xie, L.; Liang, Y.; Hong, G.; Dai, H. *J. Am. Chem. Soc.* **2011**, 133(19), 7296-7299. Liao, L.; Zhu, J.; Bian, X.; Zhu, L.; Scanlon, M. D.; Girault, H. H.; Liu, B. *Adv. Funct. Mater.* **2013**, 23, 5326-5333. Bonde, J.; Moses, P. G.; Jaramillo, T. F.; Nørskov, J. K.; Chorkendorff, I. *Faraday Discuss.* **2008**, 140, 219-231. Merki, D.; Vrabel, H.; Rovelli, L.; Fierro, S.; Hu, X. *Chem. Sci.* **2012**, 3, 2515.
5. Appel, A. M.; Lee, S.-J.; Franz, J. A.; DuBois, D. L.; DuBois, M. R. *J. Am. Chem. Soc.* **2009**, 131, 5224-5232. Karunadasa, H. I.; Montalvo, E.; Sun, Y.; Majda, M.; Long, J. R.; Chang, C. J.; *Science* **2012**, 335, 698.
6. Zhao, Y.; Truhlar, D. G. *J. Chem. Phys.*, **2006**, 125, 194101.
7. Zhao, Y.; Truhlar, D. G. *Theor. Chem. Account.* **2008**, 120, 215-241.
8. Melius, C. F.; Olafson, D. B.; Goddard, W. A., III. *Chem. Phys. Letters* **1974**, 28, 457-462. Melius, C. F.; Goddard, W. A., III. *Phys. Rev. A* **1974**, 10, 1528.
9. Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, 82, 270. Wadt, W. R.; Hay, P. J. *J. Chem. Phys.* **1985**, 82, 284.
10. Hehre, W. J.; Ditchfield, R.; Pople, J. A. *J. Chem. Phys.* **1972**, 56, 2257. Francl, M. M.; Pietro, W. J., Hehre, W. J.; Binkley, J. S.; Gordon, M. S.; DeFrees, D. J.; Pople, J. A. *J. Chem. Phys.* **1982**, 77, 3654. Hariharan, P. C.; Pople, J. A. *Theoretica Chimica Act.* **1973**, 28, 213-222
11. Krishnan, R.; Binkley, J. S.; Seeger, R.; Pople, J. A. *J. Chem. Phys.* **1980**, 72, 650.
12. Clark, T.; Chandrasekhar, J.; Spitznagel, G. W.; Schleyer, P. V. R. *Journal of Computational Chemistry*, **1983**, 4, 294-301.

13. McLean, A. D.; Chandler, G. S. *J. Chem. Phys.* **1980**, 72, 5639. Frisch, M. J.; Pople, J. A.; Binkley, J. S. *J. Chem. Phys.* **1984**, 80, 3265.
14. Tannor, D. J.; Marten, B.; Murphy, R.; Friesner, R. A.; Sitkoff, D.; Nicholls, A.; Honig, B.; Ringnalda, M.; Goddard, W. A. III. *J. Am. Chem. Soc.* **1994**, 116, 11875-11882.
15. Bochevarov, A. D.; Harder, E.; Hughes, T. F.; Greenwood, J. R.; Braden, D. A.; Philipp, D. M.; Rinaldo, D.; Halls, M. D.; Zhang, J.; Friesner, R. A. *Int. J. Quantum Chem.* **2013**, 113, 2110-2142.
16. Tissandier, M. D.; Cowen, K. A.; Feng, W. Y.; Gundlach, E.; Cohen, M. H.; Earhart, A. D.; Coe, J. V.; Tuttle, T. R., Jr. *J. Phys. Chem. A* **1998**, 102, 7787-7794.
17. Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, 77, 3865. Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1997**, 78, 1396.
18. Blöchl, P. E.; *Phys. Rev. B* **1995**, 50, 17953. Kresse, G.; Joubert, D. *Phys. Rev. B* **1999**, 59, 1758.
19. Kresse, G.; Hafner, J. *Phys. Rev. B* **1993**, 47, 558. Kresse, G.; Hafner, J. *Phys. Rev. B* **1994**, 49, 14351. Kresse, G.; Furthmüller, J. *Comput. Mat. Sci.* **1996**, 6, 15. Kresse, G.; Furthmüller, J. *Phys. Rev. B* **1996**, 54, 11169.
20. Wang, Q. H.; Kalantar-Zadeh, K.; Kis, A.; Coleman, J. N.; Strano, M. S. *Nature Nanotechnology*, **2012**, 7, 699-712.
21. Wang, H.; Zhang, Q.; Yao, H.; Liang, Z.; Lee, H.-W.; Hsu, P.-C.; Zheng, G.; Cui, Y. *Nano Lett.* **2014**, 14, 7138-7144.
22. Yu, Y.; Huang, S.-Y.; Li, Y.; Steinmann, S. N.; Yang, W.; Cao, L. *Nano Lett.* **2014**, 14(2), 553-558.

23. Helveg, S.; Lauritsen, J. V.; Laegsgaard, E.; Stensgaard, I.; Nørskov, J. K.; Clausen, B. S.; Topsøe, H.; Besenbacher F. *Phys. Rev. Lett.* **2000**, 84, 951-954.
24. Lauritsen, J. V.; Bollinger, M. V.; Laegsgaard, E.; Jacobsen, K. W.; Nørskov, J. K.; Clausen, B. S.; Topsøe, H.; Besenbacher, F. *J. Catal.* **2004**, 221, 510-522.
25. Schweiger, H.; Raybaud, R.; Kresse, G.; Toulhoat, H. *J. Catal.* **2002**, 207, 76-87.
26. Moses, P.G.; Hinnemann, B.; Topsøe H.; Nørskov, J. K. *J. Catal.* **2007**, 248, 188-203
27. Hansen, L. P.; Ramasse, Q. M.; Kisielowski, C.; Brorson, M.; Johnson, E.; Topsøe, H.; Helveg, S. *Angew. Chem. Int. Ed.* **2011**, 50, 10153-10156.
28. Li, T.; Galli, G. *J. Phys. Chem. C* **2007**, 111, 16192-16196.
29. Jaramillo-Botero, A.; Nielsen, R.; Abrol, R.; Su, J.; Pascal, T.; Mueller, J.; Goddard, W. A. III. *Top. Curr. Chem.* **2012**, 307, 1-42. Marenich, A. V.; Olson, R. M.; Kelley, C. P.; Cramer, C. J.; Truhlar, D. G. *J. Chem. Theory Comput.* **2007**, 3, 2011-203.
30. Prodhomme, P.-Y.; Raybaud, P.; Toulhoat, H. *J Catal.* **2011**, 280, 178-195.
31. Voiry, D.; Salehi, M.; Silva, R.; Fujita, T.; Chen, M.; Asefa, T.; Shenoy, V. B.; Eda, G.; Chhowalla, M. *Nano Lett.* **2013**, 13(12), 6222-6227.
32. Li, P.; Yu, Y.; Huang, Y.; Nielsen, R. J.; Goddard, W. A. III; Li, Y.; Cao, L. *ACS Catal.* **2015**, 5, 448-455.
33. Nørskov, J. K., Bliggard, T.; Logadottir, A.; Kitchin, J. R.; Chen, J. G.; Pandelov, S.; Stimming, U. *J. Electrochem. Soc.* **2005**, 152, J23-26. Nørskov, J. K.; Bligaard, T., Rossmeisl, J.; Christensen, C.H. *Nature Chemistry*, **2009**, 1, 37-46.

Supporting Information

S1 – Free Energy Barrier from Experimental Estimate

From [1], the experimental turnover frequency (TOF) is estimated to be $1.64 \times 10^{-2} \text{ s}^{-1}$ per edge Mo atom on the $(10\bar{1}0)$ Mo-edge. Since the MoS_2 used in this study models the 3-Mo wide segment of the edge, we have $\text{TOF} \approx (k_B T/h) \times \exp(-\Delta G^\ddagger/RT)$. Using $k_B = 1.38 \times 10^{-23} \text{ m}^2 \text{ kg s}^{-2} \text{ K}^{-1}$, $T = 298\text{K}$, $h = 6.63 \times 10^{-34} \text{ J}\cdot\text{s}$, $R = 1.987 \times 10^{-3} \text{ kcal}/(\text{K}\cdot\text{mol})$, ΔG^\ddagger is found to be 19.9 kcal/mol.

Reference

1. Jaramillo, T., Jørgensen, K., & Bonde, J. (2007). *Science*, 317(5834), 100–2.

*Chapter 3*THE REACTION MECHANISM FOR THE HYDROGEN EVOLUTION
REACTION ON THE BASAL PLANE SULFUR VACANCY SITE OF
MoS₂ USING GRAND CANONICAL POTENTIAL KINETICS**Abstract**

We develop the grand canonical potential kinetics (GCP-K) formulation based on thermodynamics from quantum mechanics calculations to provide a fundamental basis for understanding heterogeneous electrochemical reactions. Our GCP-K formulation arises naturally from minimizing the free energy using a Legendre transform relating the net charge of the system and the applied voltage. Performing this macroscopic transformation explicitly allows us to make the connection of GCP-K to the traditional Butler-Volmer kinetics. Using this GCP-K-based free energy, we show how to predict both the potential and pH dependent chemistry for a specific example, the hydrogen evolution reaction (HER) at a sulfur vacancy on the basal plane of MoS₂. We find that the rate determining steps in both acidic and basic conditions are the Volmer reaction in which the second hydrogen atom is adsorbed from the solution. Using our GCP-K formulation, we show that the stretched bond distances change continuously as a function of the applied potential. This shows that the main reason for the higher activity in basic conditions is that the transition state is closer to the product, which leads to a more favorable Tafel slope of 60mV/dec. In contrast, if the transition state were closer to the reactant, where the transfer coefficient is less than 0.5, we would obtain a Tafel slope of almost 150mV/dec. Based on this detailed understanding of the reaction mechanism, we conclude that **the second hydrogen at the chalcogenide vacant site is the most active towards the hydrogen evolution reaction**. Using this as a descriptor, we compare to the other 2H group VI metal dichalcogenides and predict that vacancies on **MoTe₂ will have the best performance towards HER**.

1. Introduction

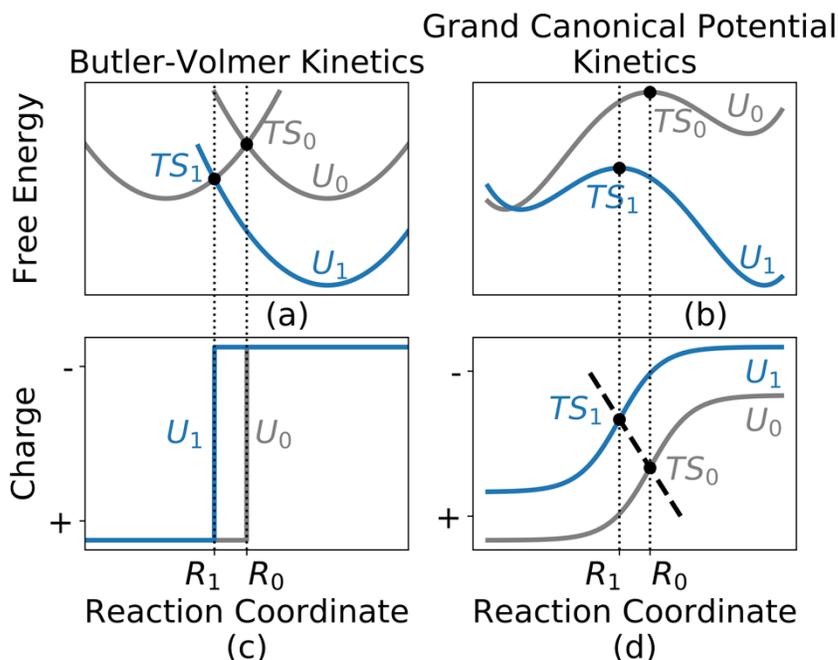


Figure 1. Schematic showing how voltage dependent electrochemical reactions described by grand canonical potential kinetics (GCP-K). (b) and (d) differs from the more standard view of Butler-Volmer kinetics (a) and (c). As the voltage is changed from U_0 to U_1 , the energy profiles shift as in (a) and (b), while the relevant reaction coordinate changes from R_0 to R_1 . The Butler-Volmer picture in (c) can be considered as a special case of the GCP-K scheme (d) in which the electron transfers instantaneously.

The field of heterogeneous electrochemistry has been growing rapidly, particularly with a focus on electrochemical water splitting and CO_2 reduction to efficiently convert electrical energy generated from traditional and renewable energy sources to recyclable energy carriers like H_2 or carbon based fuels¹⁻⁴. Simultaneously, advances in Quantum Mechanics (QM) based methods now enable the detailed reaction mechanisms to be determined for simple low index models of the catalytic surfaces⁵. Electrocatalysis is driven by applying a voltage across the reaction cell, providing a sensitive control of the rate not available with traditional heterogeneous thermal catalysis in which only the temperature and pressure can be used to

drive the chemical reactions. Recently, we have developed modifications in the traditional QM (fixed numbers of electrons) to enable the applied voltage (U) to be fixed, Grand Canonical QM (GC-QM)⁶. In GC-QM, the charges change continuously during the electrochemical reaction to keep the applied U constant. This provides a new way (Grand Canonical Potential Kinetics, GCP-K) to understand the kinetics, completely different from the traditional Butler-Volmer description of electrochemistry in which the potential surface is followed for each species keeping an integer number of electrons, from which the system can transform to a product state by tunneling between the electrode and the reacting molecule. This is illustrated in Figure 1.

Although the voltage dependent grand canonical potential can be obtained from quantum mechanical calculations, the connection to the Butler-Volmer kinetics is non-trivial as the latter theory is formulated for integer-charged solvated molecules. In this paper, we provide a macroscopic theoretical foundation for a new understanding of heterogeneous electrochemistry based on GCP-K and compare it to the traditional Butler-Volmer description.

Particularly, we will show that the voltage-dependent grand canonical potential (GCP) for surface states can be derived from traditional fixed-electron based free energies by using a Legendre transformation. As a result, we find that GCP depends quadratically on the applied potential U and on the number of electrons allowing a continuous description of the evolution of the reaction intermediates and transition states.

To illustrate the concepts underlying this new theoretical formulation, we applied the GCP-K to study the detailed reaction mechanism for the hydrogen evolution reaction (HER) on the basal plane of MoS₂. Over the past decade, many theoretical and experimental studies have shown that MoS₂ and other transition metal dichalcogenides can produce hydrogen gas efficiently⁷⁻⁸. Initially, it was believed that hydrogen atom adsorption energy on the edge sulfur (S-H bond) provided the most active site⁹⁻¹⁰. However, our QM study of the HER mechanism¹¹ showed that the rate determining step (RDS) for dihydrogen formation at edge sites takes place via the Heyrovsky reaction, in which a hydrogen (proton) from the solution

(H_3O^+) reacts with the Mo–H metal hydride bond at the edge to form H_2 . The S–H bonded site is not a kinetically important intermediate.

In addition to the edges of MoS_2 , the HER has been studied extensively for other reactive sites. This includes the 1T phase of MoS_2 ^{12,13} and amorphous MoS_x ¹⁴⁻¹⁷. Activating the basal plane of MoS_2 is of interest because of the potentially abundant number of active sites. Early studies showed that the untreated basal plane performs some HER catalysis, but the performance is less favorable than the edge sites¹⁸⁻²¹. Later, it was found that artificially creating sulfur vacancies using Ar plasma showed a correlation between the number of vacant sites and HER activity²². However, the experimental conclusion of this study is controversial, since others showed that the creation of sulfur vacancy by plasma is *not* responsible for the HER²³. Later, it was found that under electrochemical conditions sulfur vacancies are created without the use of Ar plasma²⁴. However, the interpretation is complicated because the observed reactivity might also arise from the presence of edge sites since they are known to be active towards HER. This complicates the experimental identification of the true nature of the active sites. Thus, although many studies have been performed to optimize the basal plane for HER, the reaction mechanism is not yet established.

In this study, we performed QM calculations using our new grand canonical potential (GCP-K) formulation to determine the reaction steps as a function of applied potential involved in HER at sulfur vacancies on the basal plane of MoS_2 . By accounting for all HER related chemical processes, we predict the Tafel plots and onset potentials in both the acidic and basic conditions.

Particularly, we resolved the difference in activity between the acidic and basic conditions²⁵. We focus the detailed discussions on MoS_2 , but we report the results for these new methods applied to the other transition metal dichalcogenides having the 2H structure, predicting that MoTe_2 is the best and MoS_2 is the worst for HER via basal plane vacancies.

In contrast to the simplistic view in which the performance descriptor is protonation of the reactive site, we find that it is the addition of the second hydrogen to the reactive site that is the important step. Indeed, this Volmer step determines the reaction rates in both acidic and basic conditions. Thus, we conclude that **the adsorption energy of the second hydrogen atom can be used as the proper descriptor** to assess performance for the class of group VI transition metal dichalcogenides.

2. The Grand Canonical Potential (GCP) formulation using the constant charge condition

Quantum mechanics (QM) calculations, such as density functional theory (DFT), are nearly always performed with a **fixed number of electrons**. To appropriately account for electrochemical conditions at a specified applied voltage, we must modify the methodology for the QM. Early methods to correct the QM for electrochemical systems obtained a relationship between the number of electrons and the work function of the slab surface, where surface coverage²⁶, explicit ions, or uniform background charges^{27,28} were introduced to modulate the work function of the system. Later, it was found that counter ions can be included in the implicit solvation model to provide a combined solvent-slab free energy, where the corresponding grand canonical potential is defined as in (1) so that electrochemistry processes can be obtained directly using $G(n; U)$ ²⁹⁻³¹.

$$G(n; U) = F(n) - ne(U_{SHE} - U) \quad (1)$$

where G is the grand canonical free energy, which depends on the applied voltage U vs. SHE, n is the number of electrons, e is unit electron volt in energy, F is the total free energy as a function of n , and $U_{SHE} = \mu_{e,SHE}/e$ is the electronic energy at the standard hydrogen electrode (SHE) condition. The signs are chosen such that U is directly related to the experimentally defined value, i.e., $U = -0.1V$ corresponds to $-0.1V$ vs. SHE. Changing to the reference hydrogen electrode (RHE) shifts the reference fermi level further depending on the pH of the solution.

However, for $G(n; U)$ to be used as a thermodynamic potential, the number of electrons in the system must be equilibrated to the applied voltage. To do this, the QM self-consistent approach is to match the electronic fermi level to that of the applied potential by changing the occupation of the electronic bands, thus varying the number of electrons³²⁻³³. Mathematically, this is

$$\mu_e = \frac{dF(n)}{dn} = e(U_{SHE} - U) \text{ or } \frac{dG(n;U)}{dn} = 0 \quad (2)$$

Thus, we *define* the macroscopic thermodynamic **Grand Canonical Potential** (GCP) as in (3).

$$\text{GCP}(U) = \min_n G(n; U) = \min_n (F(n) - ne(U_{SHE} - U)) \quad (3)$$

Since experimental observations typically involve the response of a chemical system as a function of the applied voltage, we recommend using $\text{GCP}(U)$ directly as an explicit function of U in QM calculations modeling electrochemical processes. In contrast, many recent studies *have assumed* a $\text{GCP}(U)$ that *depends linearly on* U ³⁴⁻³⁵.

Instead, our definition of the $\text{GCP}(U)$ in the form of minimization as in Equation (3), makes it immediately obvious that *the linear approximation is not correct*. **The form of $F(n)$ must be at least quadratic in n** in order to describe the minimization of $\text{GCP}(n; U)$. As reported previously²⁷⁻²⁸, the form of $\text{GCP}(U)$ is in fact approximately quadratic. Hence, we expand $F(n)$ in a quadratic form

$$F(n) = a(n - n_0)^2 + b(n - n_0) + c,$$

where a , b , and c are fitted parameters. Substituting and performing the minimization, we have

$$\text{GCP}(U) = -\frac{1}{4a} (b - \mu_{e,SHE} + eU)^2 + c - n_0 \mu_{e,SHE} + n_0 eU \quad (4)$$

Using this form, we relate the parameters a , b , and c to physical quantities as follows:

- First, when the system is neutral, $F(n = n_0) = c$.
- Second, the number of electrons is $n(U) = -\frac{1}{e} \frac{\partial GCP(U)}{\partial U} = n_0 - \frac{1}{2ae} (b - \mu_{e,SHE} + eU)$.
Thus at the potential of zero charge, $n(U_{PZC})=n_0$, and we obtain $b = \mu_{e,SHE} - eU_{PZC}$.
- Finally, the differential capacitance is $C_{diff} = \frac{\partial n}{\partial U} = -\frac{1}{2a}$, which gives $a = -\frac{1}{2C_{diff}}$.

Summarizing, the grand canonical potential and the free energy have the following form in terms of physical quantities:

$$GCP(U) = \frac{e^2 C_{diff}}{2} (U - U_{PZC})^2 + n_0 eU + F_0 - n_0 \mu_{e,SHE} \quad (5a)$$

$$F(n) = -\frac{1}{2C_{diff}} (n - n_0)^2 + (\mu_{e,SHE} - eU_{PZC})(n - n_0) + F_0 \quad (5b)$$

where

- C_{diff} is the differential capacitance, calculated from parameter a ,
- U_{PZC} is the potential of zero net charge, calculated from parameter b ,
- F_0 is the free energy at zero net charge, calculated from parameter c ,
- n_0 is the number of electrons at zero net charge, summing all valence electrons in the QM,
- $\mu_{e,SHE}$ is the chemical potential of an electron vs. SHE, and
- e is the energy of an electron volt, which is for unit conversion from voltage to energy

The quadratic dependence in the free energy $F(n)$ and the grand canonical potential $GCP(U)$ implies that the capacitive effects will participate in the electrochemical processes. Fundamentally, this is due to the fact that heterogeneous systems allow electrons to

delocalize into broad electronic bands, resulting in fractional occupations. Thus, the number of electrons can vary continuously, leading to the capacitive contributions.

We note that a similar quadratic form

$$G(U) = F(n(U)) - ne(U_{\text{SHE}} - U)$$

was proposed previously³⁶, but the relationship between n and U was established via the work function rather than a proper thermodynamic minimization.

Our approach shows that the work function is not needed to calculate the grand canonical potential since we define $GCP(U)$ rigorously from the free energy $F(n)$ via a Legendre transform. Using the Legendre transform allows us to write $F(n)$ and $GCP(U)$ in terms of physical parameters, with the connection to Butler-Volmer kinetics as discussed below.

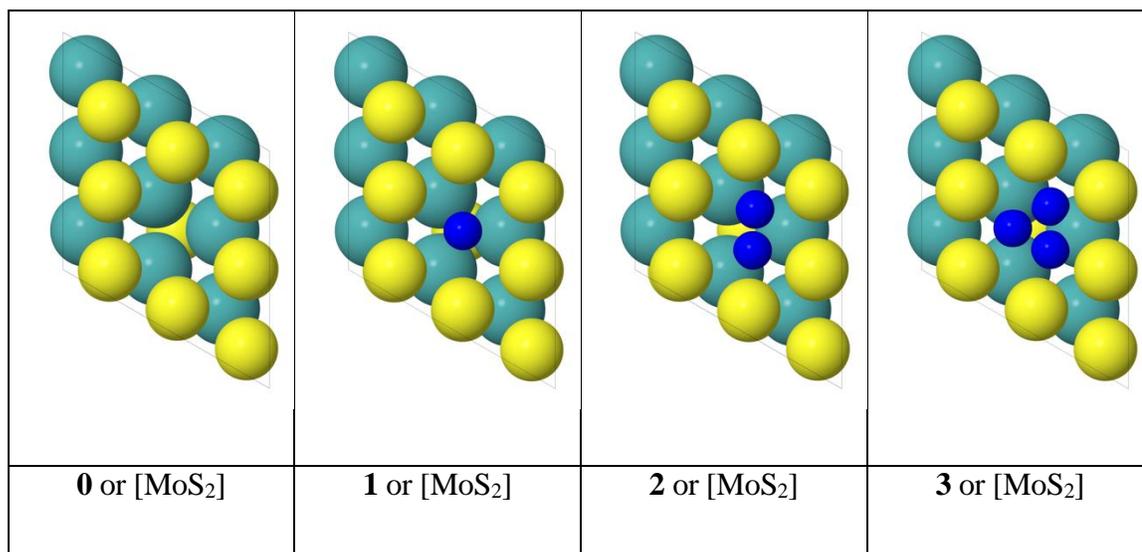


Figure 2. Four possible states for Hydrogen adsorption at the sulfur vacancy site. Blue: hydrogen atom; yellow: sulfur atom; cyan: molybdenum atom.

3. Simulation Model for the Basal Plane of MoS₂

To predict the QM properties of a sulfur vacancy on the basal plane of MoS₂, we used a 3x3 MoS₂ periodic slab. Removing a sulfur atom exposes three molybdenum atoms that become available for bonding. This leads to four possible intermediate states with 0 to 3 hydrogen

atoms at the vacant site. We label them as **0**, **1**, **2** and **3** or [MoS₂], [MoS₂]H, [MoS₂]H₂ and [MoS₂]H₃ in Figure 2. A detailed description of the quantum mechanical calculations is provided in the supplementary information.

To illustrate the quadratic behavior of the grand canonical potential, we examined the voltage dependence for [MoS₂]H in detail. [MoS₂]H is used here because we find below that it is both the most stable intermediate thermodynamically and most populous for the steady state reaction. Figure 3(a) shows that the solvated free energy $F(n)$ as a function of accumulated charges *appears* to be linear. However, this is due to the large contribution from the free energy of an electron at SHE. Rearranging Equation (5), leads to (6),

$$F(n) - \mu_{e,SHE}n = -\frac{1}{2C_{diff}}(n - n_0)^2 - \mu_{e,SHE}n_0 - eU_{PZC}(n - n_0) + F_0 \quad (6)$$

Equation (6) reveals the quadratic dependence on charge, as shown in Figure 3(b). We see that the minimum is approximately at $n=n_0$.

Under electrochemical conditions, a nonzero voltage, U , is applied to drive the reaction. This shifts the free energy in Equation (6) by neU , leading to $G(n, U)$, as defined in Equation (1). Figure 3(c) shows that applying a voltage of $U = -0.5V$ vs. SHE shifts the minimum of $G(n; U)$ towards more electrons, indicating that the slab becomes more negatively charged.

Thus, although the free energy $F(n)$ appears to be linear in Figure 3(a), the thermodynamically relevant potential $G(n; U)$ or GCP(U) in Figure 3(c) is clearly quadratic.

4. Relationship between grand canonical potential reaction kinetics and Butler-Volmer reaction kinetics

The above formulation can be used for stable states because the equilibrium geometries usually change little as the applied potential changes. However, since the transition state (TS) is the maximum along the *minimum energy path* (MEP) for the reaction coordinate, *the TS geometry will change as the applied voltage changes.*

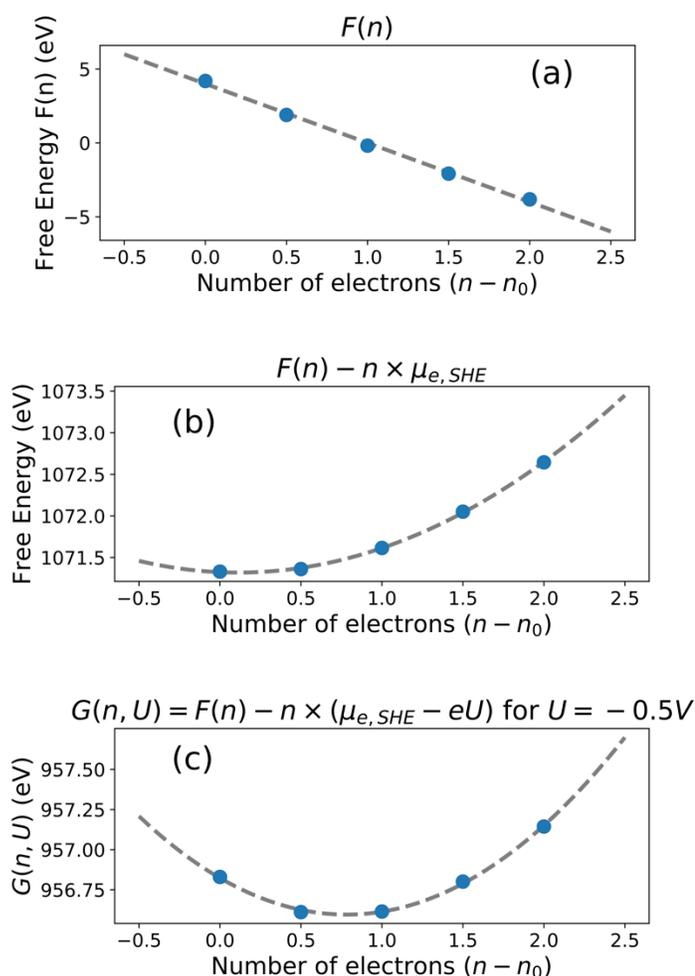


Figure 3. The free energy and grand canonical potential as a function of the number of electrons. The DFT energies are indicated by blue dots, the dashed curve is the polynomial fit. (a) a linear fit to $F(n)$, (b) a quadratic fit to $F(n) - n \times \mu_{e,SHE}$, (c) a quadratic fit GCP(n, U).

Fig 1 (b) shows schematically that as the voltage is changed from U_0 to a more negative U_1 , the negatively charged product becomes more stable, shifting the potential energy surface downward for the species. However, this energy changes as the geometries change along the MEP. Because the reactant has fewer electrons, the stabilization is less effective, resulting in a leftward shift of the transition state towards the reactant. The coordinate along the reaction coordinate changes from R_0 to R_1 in Figure 1(b). In comparison, for Butler-Volmer kinetics³⁷⁻³⁹, only two states are involved, and the shift from R_0 to R_1 is the result of the shift in energy for the final state, as shown in Figure 1(a).

The difference between the *grand canonical potential reaction kinetics* (GCP-K) and Butler-Volmer kinetics is more obvious for the reaction path in the charge-reaction coordinate or (n, R) plane as shown in Figure 1(c) and 1(d). In the Butler-Volmer picture, an electron is transferred through tunneling from the electrode to the product, resulting in a discontinuity in the (n, R) plane. However, in extended systems where intermediates are adsorbed on the surface, there can be fractional charges per unit area since electrons are delocalized. As a result, the surface species can charge or discharge continuously, leading to a smooth reaction path in the (n, R) plane, as shown in Figure 1(d). Thus, the Butler-Volmer picture in the (n, R) plane can be considered as the special case of the GCP-K picture in which the electron transfer takes place *instantaneously* as in Figure 1(c).

In the GCP-K picture, the reaction path changes continuously in the (n, R) plane. Thus, both the *charge* n and the *spatial reaction coordinate* R are relevant coordinates. Because the constant charge free energy $F(n)$ is used to transform to $GCP(U)$, we must prove that the grand canonical potential for the transition state obtained from the constant charge $F(n)$ coincides with the grand canonical potential obtained from constant voltage calculations.

The transition state grand canonical potential $GCP_{TS,n}(U)$ can be found explicitly by transforming $F_{TS}(n)$, where $F_{TS}(n)$ is the barrier for each fixed charge n such that

$$F_{TS}(n) = \max_R F(n, R), \text{ with } R_{IS} < R < R_{FS}$$

Then,

$$GCP_{TS,n}(U) = \min_R (F_{TS}(n) - ne(U_{SHE} - U))$$

On the other hand, including the spatial dependence in Equation (1), leads to

$$G(n, R; U) = F(n, R) - ne(U_{SHE} - U)$$

Thus, the barrier calculated from the explicit voltage dependent grand canonical potential is defined as

$$\text{GCP}_{\text{TS},U}(U) = \max_R \text{GCP}(U, R) = \max_R \min_n \text{GCP}(n, R; U)$$

To show that the two approaches, $\text{GCP}_{\text{TS},U}(U)$ and $\text{GCP}_{\text{TS},n}(U)$, are equivalent, we employ the minimax theorem⁴⁰:

Given:

1. $F(n, R)$ is quadratic and thus convex in n , then so is $\text{GCP}(n, R; U) = F(n, R) - ne(U_{\text{SHE}} - U)$,
2. The reaction path is smooth in extended systems since the charges transfer continuously at the electrode. By the definition of the transition state, the reaction path is concave in R in the neighborhood of R_{TS} .

Then,

$$\begin{aligned} \text{GCP}_{\text{TS},U}(U) &= \max_R \min_n (\text{GCP}(n, R; U)) \\ &= \min_n \max_R (\text{GCP}(n, R; U)) \\ &= \min_n \max_R (F(n, R) - ne(U_{\text{SHE}} - U)) \\ &= \min_n (F_{\text{TS}}(n) - ne(U_{\text{SHE}} - U)) \\ &= \text{GCP}_{\text{TS},n}(U) \end{aligned}$$

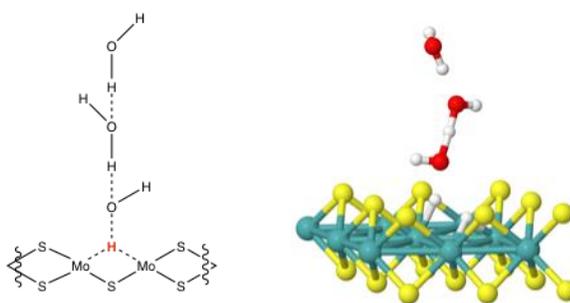
Thus, the *transition states obtained from the constant charge barriers and the constant voltage barriers are indeed equivalent*. This relationship allows us to calculate the barriers for a system with a fixed number of electrons (standard QM) and then use the Legendre transform to obtain the voltage dependence for the transition state. Figure 1(d) shows that for each voltage (U), there corresponds a transition state with a specific charge (n) and spatial distance (R_{TS}). Thus, R_{TS} is a function of charge (n), or $R_{\text{TS}}(n) = \text{argmax}_R F(n, R)$.

Because the grand canonical potential of the transition state is found by minimization in n and maximization in R , at the transition state we then have

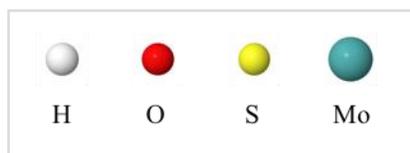
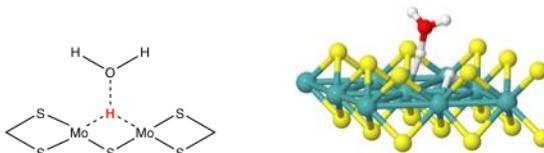
$$\frac{d}{dR} \frac{d}{dn} \text{GCP}(n, R; U) = 0 \quad (7)$$

As discussed above, $\text{GCP}(n, R; U)$ depends on R and U quadratically around the transition state, the derivatives in n and R reduce the dependence from order 2 to order 1. As a result, Equation (7) implies that $R_{TS}(n)$ depends linearly on n . This linear dependence is shown qualitatively as the dashed line in Figure 1(d).

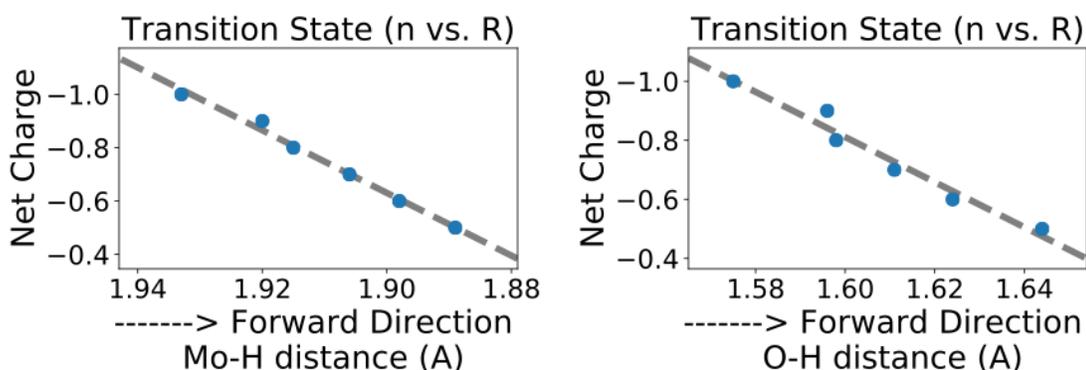
(a) Schematics of TS in base:



(b) Schematics of TS in acid:



(c): Mo-H and O-H as functions of net charges in base:



(d): Mo-H and O-H as functions of net charges in acid:

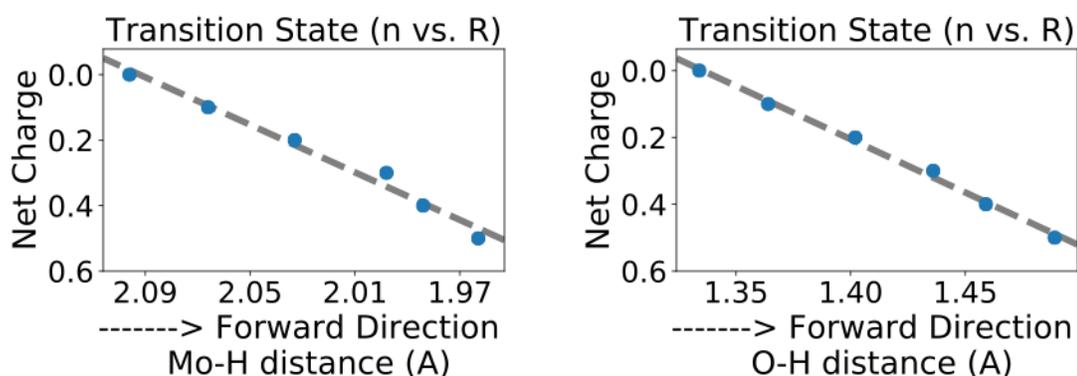


Figure 4. The transition states as the system evolves from $[\text{MoS}_2]\text{H}$ to $[\text{MoS}_2]\text{H}_2$. (a): the transition state structure in basic conditions. (b): the transition state structure in acidic conditions. (c): relationships between the transition state charge (n) and spatial coordinates (R_{TS}) in basic conditions. (d): relationships between the transition state charge (n) and spatial coordinates (R_{TS}) in acidic conditions. The linear relationships in (c) and (d) agree with the qualitative picture in Figure 1(d).

To show this linear relationship quantitatively, we studied the transition state of hydrogen transfer from the solution to $[\text{MoS}_2]\text{H}$ to produce $[\text{MoS}_2]\text{H}_2$, as shown in Figure 4. We show below that this is the rate determining step (RDS) for HER at the sulfur vacancy site on the basal plane of MoS_2 . In this primary step, the hydrogen atom gradually moves from the oxygen atom of the water molecule towards the molybdenum atom at the reaction center. As a result, two spatial coordinates are important. One is the Mo-H distance and the other is

O–H distance, such that in the forward direction, the distance of Mo–H becomes shorter and the distance of O–H is becoming longer. Figure 4 (e)-(h) shows that the relationship between R and n for the transition state is linear as predicted by the GPC-K approach.

5. Discussion of the HER mechanism on MoS₂ basal plane

5.1 The Hydrogen Evolution Reaction in Acidic Conditions

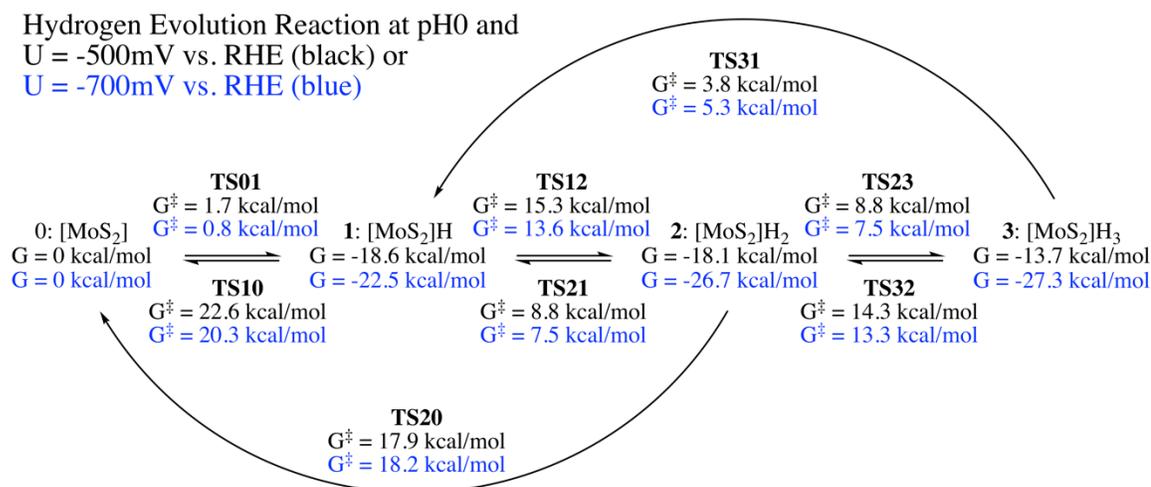


Figure 5. The free energies at 298K under acid conditions for all reaction intermediates and transition states involved in the HER at the sulfur vacancy on the basal plane of MoS₂. Black: $U = -500\text{mV}$ vs RHE, blue: $U = -700\text{mV}$ vs. RHE.

Using the information from Figure 5, we can calculate the grand canonical potential of all the relevant intermediates and their connecting transition states at any given applied potential. In this reaction, the sulfur vacancy site can bind up to three hydrogen atoms.

- At $U = -500\text{mV}$, the most stable state is **1**, or [MoS₂]H.
- However, at a more negative potential, $U = -700\text{mV}$, the most stable state becomes **2**, or [MoS₂]H₂.

This is expected, because the grand potential $G(n, U) = F(n) - neU$ is smaller for more negative charge (larger n) when a negative potential (more negative U) is applied. Thus, the [MoS₂]H₂ becomes more stable than [MoS₂]H with a more negative charge.

5.1a starts with the sulfur vacant site with no hydrogen atom adsorbed:

- At $U = -500\text{mV}$, we first adsorb a hydrogen atom from the solution at **TS01** with a barrier of 1.7 kcal/mol,
- At $U = -700\text{mV}$, we first adsorb a hydrogen atom from the solution at **TS01** with a barrier of 0.8 kcal/mol.

5.1b starts with one hydrogen atom adsorbed at the site:

This site can react with another hydrogen from the solution to *generate an H_2* , while leaving behind the empty site with

- A barrier (**TS10**) of 22.6 kcal/mol at $U = -500\text{ mV}$ and
- A barrier (**TS10**) of 20.3 kcal/mol at $U = -700\text{ mV}$,

Or, it can abstract a second hydrogen from the solution to form $[\text{MoS}_2]\text{H}_2$ with

- A barrier of 15.3 kcal/mol (**TS12**) at $U = -500\text{ mV}$ and
- A barrier of 13.6 kcal/mol (**TS12**) at $U = -700\text{ mV}$.

5.1c starts with two hydrogen atoms adsorbed at the site:

They can react with each other to *generate an H_2* while leaving behind the empty site with

- A barrier (**TS20**) of 17.9 kcal/mol at $U = -500\text{ mV}$, and
- A barrier (**TS20**) of 18.2 kcal/mol at $U = -700\text{ mV}$.

or one of the hydrogen atoms can react with another hydrogen atom from water to generate H_2 while leaving behind $[\text{MoS}_2]\text{H}$ with

- A barrier of 8.8 kcal/mol at $U = -500\text{ mV}$ (**TS21**). and
- A barrier of 7.5 kcal/mol at $U = -700\text{ mV}$ (**TS21**).

or $[\text{MoS}_2]\text{H}_2$ can abstract another hydrogen solution to form $[\text{MoS}_2]\text{H}_3$ via **TS23**. However, we found that this step involves the same transition state as **TS21**, thus the barriers are 8.8 and 7.5 kcal/mol.

For $[\text{MoS}_2]\text{H}_3$, no more hydrogen can be added, thus it can generate H_2 via the Volmer step (**TS31**) with a barrier of 3.8 and 5.3 kcal/mol, or it can generate H_2 via the Heyrovsky step (**TS32**) with a barrier of 14.3 and 13.3 kcal/mol.

5.2 Hydrogen Evolution Reaction in Basic Conditions

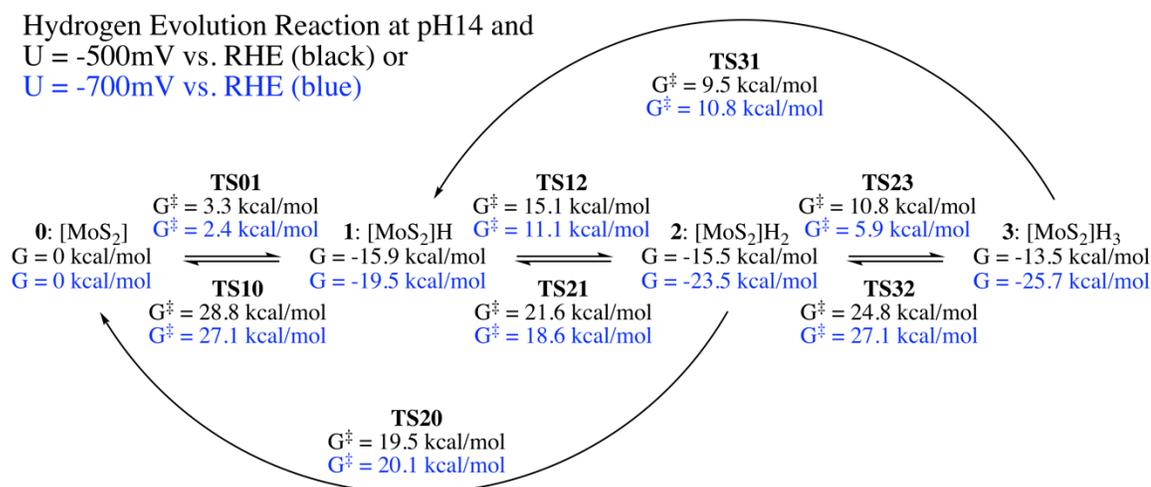


Figure 6. The free energies at 298K under basic conditions for all the reaction intermediates and transition states involved in HER at the sulfur vacancy on the basal plane of MoS₂. Black: -500mV vs RHE, blue: -700mV vs. RHE.

Similarly, we can use the GCP-K method in basic conditions to calculate the GCP energy of each species and the barrier for each reaction. However, because we use RHE, the reference fermi energy of the electron is shifted by $\text{pH} \times 0.059$ eV. As shown in Figure 6, at $U = -500\text{mV}$ vs RHE, the relative stabilities of $[\text{MoS}_2]\text{H}$, $[\text{MoS}_2]\text{H}_2$ and $[\text{MoS}_2]\text{H}_3$ are very similar, while $[\text{MoS}_2]$, the state with no hydrogen atom adsorbed, is much less stable. Thus, under basic conditions, we conclude that there is always at least a hydrogen atom adsorbed at the reaction site. Specifically,

- At $U = -500$ mV, the most stable state is **1**, or $[\text{MoS}_2]\text{H}$, the same as the acidic case, at $U = -700$ mV, the most stable state is shifted to **3**, or $[\text{MoS}_2]\text{H}_3$.

This shows that at high applied potential, there is a bias toward the intermediate species with more electrons, making them more stable.

5.2a starts with no hydrogen adsorbed:

- At $U = -500\text{mV}$, the barrier **TS01** leading to $[\text{MoS}_2]\text{H}$ is very low at 3.3 kcal/mol,
- At $U = -700\text{mV}$, **TS01** is even lower at 2.4 kcal/mol.

Such a low barrier is due to the large thermodynamic force, so that $[\text{MoS}_2]\text{H}$ is -15.9 and -19.5 kcal/mol downhill from $[\text{MoS}_2]$ for -500mV and -700mV , respectively. However, the

kinetics at $[\text{MoS}_2]\text{H}$ is more unfavorable because it is relatively too stable. It has to overcome a large thermodynamic force to react with another water molecule to form H_2 and $[\text{MoS}_2]$ via **TS10**.

5.2b starts with one hydrogen atom adsorbed:

Different from the vacant state $[\text{MoS}_2]$, the kinetics at $[\text{MoS}_2]\text{H}$ is more unfavorable because it is relatively much more stable. To react with another water molecule to form H_2 and $[\text{MoS}_2]$ via **TS10**, it has to overcome a large thermodynamic force:

- At $U = -500\text{mV}$, the barrier **TS10** leading to $[\text{MoS}_2]$ is 28.8 kcal/mol,
- At $U = -700\text{mV}$, the barrier **TS10** is 27.1 kcal/mol.

On the other hand, it is much easier to form two adsorbed hydrogen atoms at the reaction site (**TS12**), since $[\text{MoS}_2]\text{H}_2$ is similar in energy to $[\text{MoS}_2]\text{H}$.

- A barrier (**TS12**) of 15.1 kcal/mol at $U = -500\text{mV}$,
- A barrier (**TS12**) of 13.6 kcal/mol at $U = -700\text{mV}$.

5.2c starts with two hydrogen atoms adsorbed:

Having two adsorbed hydrogen atoms for $[\text{MoS}_2]\text{H}_2$ allows the H_2 molecule to be formed via the Heyrovsky mechanism (**TS21**) or the Tafel mechanism (**TS20**). However, both steps have to overcome a $\sim 20\text{kcal/mol}$ barrier at -500mV and -700mV .

- At $U = -500\text{mV}$, the Tafel barrier (**TS20**) is 19.5 kcal/mol and the Heyrovsky barrier (**TS21**) is 21.6 kcal/mol,
- At $U = -700\text{mV}$, the Tafel barrier (**TS20**) is 20.1 kcal/mol and the Heyrovsky barrier (**TS21**) is 18.6 kcal/mol.

Instead of forming H_2 directly, $[\text{MoS}_2]\text{H}_2$ prefers to abstract one more hydrogen atom from the solution,

- A barrier (**TS23**) of 10.8 kcal/mol at $U = -500\text{mV}$,
- A barrier (**TS23**) of 5.9 kcal/mol at $U = -700\text{mV}$.

5.2d starts with three hydrogen atoms adsorbed:

At last, H_2 molecule can be formed from $[\text{MoS}_2]\text{H}_3$ via the Heyrovsky step (**TS32**) or the Tafel step (**TS31**). However, the Tafel step wins since the barrier is only 9.5 kcal/mol (-500mV) and 10.8 kcal/mol (-700mV), much lower than the activation energy for the Heyrovsky step with a barrier of almost 25 kcal/mol.

To summarize, our analysis shows that hydrogen formation at the sulfur vacant site in basic conditions starts from $[\text{MoS}_2]\text{H}$, and continues to bind two hydrogen atoms sequentially, leading to $[\text{MoS}_2]\text{H}_3$. Then H_2 molecule is formed via the Tafel mechanism while $[\text{MoS}_2]\text{H}_3$ returns to the initial state $[\text{MoS}_2]\text{H}$.

6. Overall kinetics

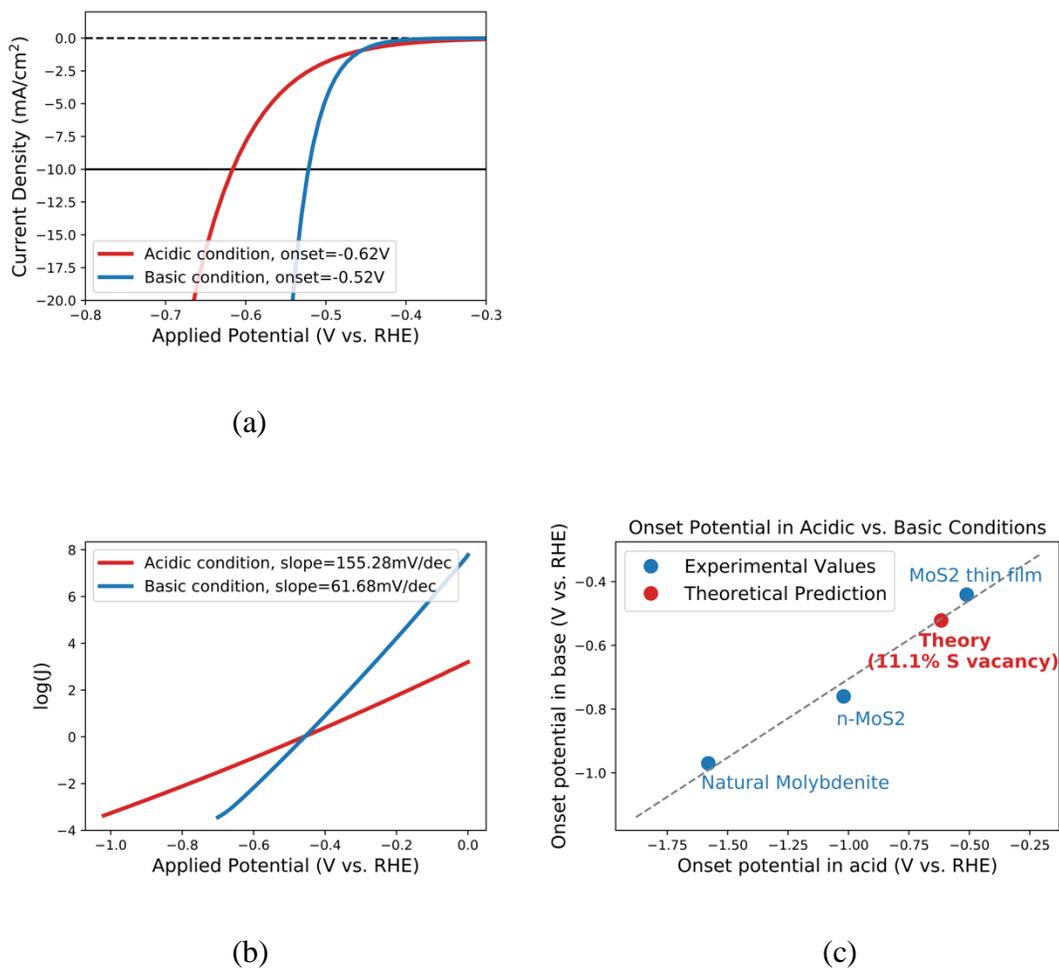


Figure 7. (a) QM predicted I-V curves for acidic and basic conditions. (b) Tafel plots for acidic and basic conditions. (c) Comparison between our QM predictions and the experimental interpolation from [A24]. This shows excellent agreement.

Since all the energies of the relevant reaction intermediates and transition states are calculated as functions of applied potential using the quadratic grand canonical potential, the

potential dependent rate constants are obtained using the Eyring rate equation. Once all rate constants are found, a microkinetic model is used to calculate the overall reaction rates and species concentrations.

First, we write the rate equations for all species as:

$$\begin{aligned}\frac{dx_0}{dt} &= -k_{01}x_0 + k_{10}x_1 + k_{20}x_2 \\ \frac{dx_1}{dt} &= k_{01}x_0 - k_{10}x_1 - k_{12}x_1 + k_{21}x_2 \\ \frac{dx_2}{dt} &= k_{12}x_1 - k_{21}x_2 - k_{20}x_2 + k_{32}x_3 \\ \frac{dx_3}{dt} &= k_{23}x_2 - k_{32}x_3 - k_{31}x_3\end{aligned}$$

and the Eyring rate equation as:

$$k_{ij}(U) = \frac{k_B T}{h} \exp\left(-\frac{\Delta G_{ij}^\ddagger(U)}{k_B T}\right)$$

where x_i is the concentration for each intermediate species, $k_{ij}(U)$ is the voltage dependent rate constant. However, the above set of equations is linearly dependent. We must include an additional constraint, $\sum_i x_i = 1$. If we are concerned only with the steady state chemistry, we can set the left hand sides of the above equation to zero. We then obtain the corresponding rates and concentrations by solving the system of linear equations. This then leads to the I-V plot and the Tafel plot as shown in Figure 7 (a) and (b).

Typically, experimental studies report the applied voltage at 10mA/cm² as the onset potential for the catalyst. As shown in Figure 7(a), we found that the onset potential for the 11.1% sulfur vacancy to Mo atom ratio is -0.62V in acidic conditions, and -0.52V in basic conditions, agreeing qualitatively with experimental findings that the basal plane of MoS₂ is more active in basic conditions than in acidic conditions.

To correlate directly with experimental onset potentials, the experimental number density of the sulfur vacancies must be known. However, if the same sites are responsible for HER in

both acidic and basic conditions, the onset potentials should scale linearly between acid and base. Indeed, this exact linear relationship is found for all reported onset potentials for various samples²⁵, which are plotted in Figure 7(c). Our predicted onset potentials for base and acid are also plotted in Figure 7(c), which agree very well with our fitted line across the experimental samples. The Tafel slope in acidic conditions is reported to be 127mV/dec, corresponding to a transfer coefficient of 0.47, which is similar to our predicted transfer coefficient of 0.39 derived from our Tafel slope of 155mV/dec in Figure 7(b). No Tafel slope has been reported for HER on the basal plane of MoS₂. We predict the Tafel slope to be 62 mV/dec.

Condition	[MoS ₂]	[MoS ₂]H	[MoS ₂]H ₂	[MoS ₂]H ₃
Acidic (pH0)	0.0(-500mV)	0.9999	7.976×10^{-6}	1.119×10^{-9}
	0.0(-700mV)	0.9998	1.480×10^{-5}	2.544×10^{-7}
Basic (pH14)	0.0(-500mV)	0.9992	6.791×10^{-4}	7.323×10^{-5}
	0.0(-700mV)	0.6542	1.100×10^{-4}	0.3457

Table 1. Predicted species concentrations in fractions at the sulfur vacancy during hydrogen evolution reaction. The concentrations are normalized to sum to 1.

As shown in Table 1, the dominant reaction intermediate is [MoS₂]H, the reaction site with one hydrogen atom adsorbed, agreeing with our discussions of Figure 5 and 6 where we conclude that [MoS₂]H is the starting point of the catalytic cycle. Although the energetics of [MoS₂]H is nearly identical to [MoS₂]H₂, as shown in Figure 5 and 6, the near unity concentration is mainly the result of the rate determining step in which [MoS₂]H is protonated to [MoS₂]H₂.

However, in basic condition, when the applied voltage is high, e.g., -700mV vs. RHE, the concentration of [MoS₂]H₃ increases to a nontrivial amount. This is mainly because the activation energy of the Tafel reaction (**TS31**) becomes e nearly as high as the original Volmer rate determining step of **TS12**.

7. Comparison with other Group VI Transition Metal Dichalcogenides

Since the adsorption of the second hydrogen atom to the sulfur vacant site is the general rate determining step, we can use its adsorption grand potential as the descriptor to compare the hydrogen evolution reaction activity at the basal plane across the class of group VI transition metal chalcogenides. The conventional adsorption energy is calculated at neutral charge, but since the number of electrons is not equilibrated to an applied voltage, this approach corresponds to intermediate states having different voltage. Such an approach is inappropriate for electrochemistry because the energy difference between two arbitrary voltages does not have any physical or chemical significance.

Instead, we compare the GCP of the intermediate species at the same voltage. We have also used the GCP approach to predict successfully the hydrogen adsorption energies at the same voltage ($V=0$ vs. SHE) to explain the hydrogen evolution relationships on MoSSe/NiSe₂⁴¹ and FeP/NiP⁴².

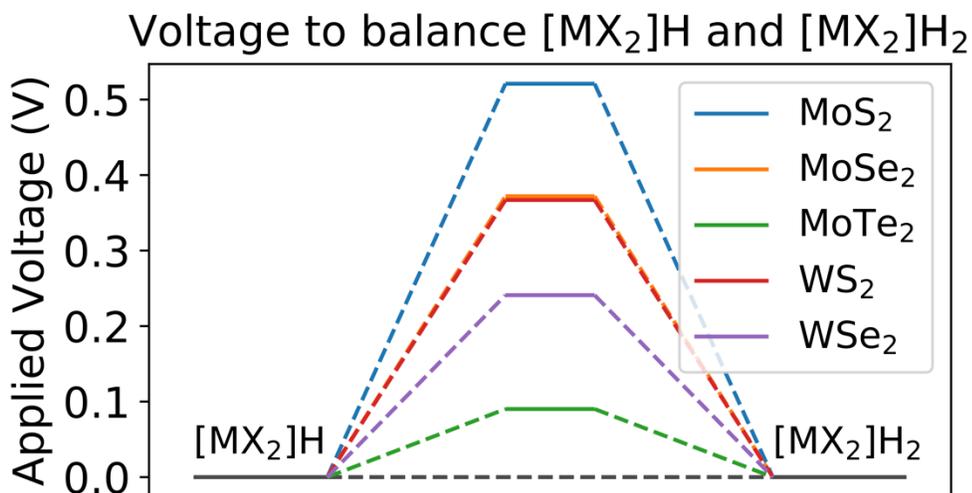


Figure 8. Required applied potential to obtain a zero reaction energy for the rate determining Volmer step from $[MX_2]H$ to $[MX_2]H_2$. We predict that this lowest required voltage will correspond to the best HER performance.

Since we have shown that the rate determining step is the Volmer reaction from $[\text{MoS}_2]\text{H}$ to $[\text{MoS}_2]\text{H}_2$, we can use the reaction energy of this step to compare the reactivity across materials with similar structures. However, the reaction energy, $\Delta G_{\text{H}_2}(U) = G_2(U) - G_1(U)$ between **2** and **1**, depends on voltage. Instead of comparing $\Delta G_{\text{H}_2}(U)$, we calculate the voltage necessary to obtain $\Delta G_{\text{H}_2}(U)=0$, then the material with the lowest required voltage will be the most active towards HER. The corresponding voltages for different transition metal dichalcogenides are reported in the first numerical column in Table 2.

	Voltage when $G([\text{MX}_2]\text{H}) =$ $G([\text{MX}_2]\text{H}_2)$	$\Delta G([\text{MX}_2])$	$\Delta G([\text{MX}_2]\text{H}_3)$
MoS ₂	$\eta = 0.521\text{V}$	0.82 eV	0.17 eV
MoSe ₂	$\eta = 0.372\text{V}$	0.83 eV	0.065 eV
MoTe ₂	$\eta = 0.0902\text{V}$	0.69 eV	0.046 eV
WS ₂	$\eta = 0.367\text{V}$	0.61 eV	0.26 eV
WSe ₂	$\eta = 0.241\text{V}$	0.65 eV	0.21 eV

Table 2. The relative grand canonical energies of $[\text{MX}_2]$ and $[\text{MX}_2]\text{H}_3$ at the optimal voltages. The relative energies $\Delta G([\text{MX}_2])$ and $\Delta G([\text{MX}_2]\text{H}_3)$ referenced to $[\text{MX}_2]\text{H}$ (or $[\text{MX}_2]\text{H}_2$) indicate whether the species might be important for the actual reaction mechanism.

In addition to the required potentials, we also calculated the relative energies of $[\text{MX}_2]$ and $[\text{MX}_2]\text{H}_3$ to determine whether they will interfere with the proposed stable states of $[\text{MX}_2]\text{H}$. As shown in Table 2, all of the relative energies of $[\text{MX}_2]$ are greater than 0.6 eV from the $[\text{MX}_2]\text{H}$ species, reaffirming that the vacant site $[\text{MX}_2]$ is not an important intermediate during HER at the basal plane. On the other hand, the $[\text{MX}_2]\text{H}_3$ are relatively much more stable, which allows the HER to proceed and complete the catalytic cycle from $[\text{MX}_2]\text{H}_3$ back to the original state of $[\text{MX}_2]\text{H}$.

Based on these calculations, we predict that MoTe₂ will have the best per site activity across the stable 2H group VI metal dichalcogenides, with $\eta=0.09$ V. Next is WSe₂ with $\eta=0.24\text{V}$, followed by WS₂ and MoSe₂ with $\eta=0.37\text{V}$, with MoS₂ last at $\eta=0.52\text{V}$. Our predicted trend agrees with the observed trend²⁵ that for single crystal MoS₂ and MoSe₂, the onset potential for MoSe₂ is -0.78V vs. RHE in acidic condition, which is 0.47V lower than the onset

potential of -1.25V vs. RHE for MoS_2 . If a transfer coefficient of 0.5 is assumed, the onset potential of MoTe_2 will be 0.22V less negative than MoS_2 for the same density of chalcogenide vacancies.

8. Conclusion

In conclusion, we have shown that our formulation of the grand canonical potential kinetics (GCP-K) in terms of thermodynamics provides a fundamental basis for understanding electrochemical processes. Our GCP-K formulation arises naturally from minimizing the free energy using a Legendre transform. As the result, the free energies and the grand canonical potentials of the reaction intermediates include a quadratic term that depends on the differential capacitance C_{diff} . We use the minimax theorem to show that the barriers in the constant charge picture and constant potential picture describe the same transition states. Using this GCP-K-based free energy, we showed how to predict both the potential and pH dependent chemistry of the hydrogen evolution reaction at the sulfur vacancy of the basal plane of MoS_2 .

We find that the rate determining steps in both the acidic and basic are the Volmer reaction in which the second hydrogen forming is adsorbed from the solution. Using our GCP-K formulation, we show that the stretched bond distances change continuously as a function of the applied potential. This shows that the main reason for the higher activity in basic conditions is that the transition state is closer to the product, leading to the much more favorable Tafel slope of $60\text{mV}/\text{dec}$. In contrast if the transition state were closer to the reactant, where the transfer coefficient is less than 0.5 we would obtain a Tafel slope of almost $150\text{mV}/\text{dec}$.

Based on this detailed understanding of the reaction mechanism, we conclude that **the second hydrogen at the chalcogenide vacant site is the most active towards the hydrogen evolution reaction**. Using this as a descriptor, we compared the rest of the 2H group VI metal dichalcogenides and predict that **MoTe_2 will have the best performance**

towards HER among the 2H group VI transition metal dichalcogenides considered here.

References

1. Walter, M. G., Warren, E. L., McKone, J. R., Boettcher, S. W., Mi, Q., Santori, E. A., Lewis, N. S. Solar water splitting cells. *Chemical Reviews*, **2010**. 110(11), 6446–6473.
2. Roger, I., Shipman, M. A., Symes, M. D. Earth-abundant catalysts for electrochemical and photoelectrochemical water splitting. *Nature Reviews Chemistry*, **2017**. Volume 1, Article number: 0003.
3. Appel, A. M., Bercaw, J. E., Bocarsly, A. B., Dobbek, H., DuBois, D. L., Dupuis, M., Waldrop, G. L. Frontiers, Opportunities, and Challenges in Biochemical and Chemical Catalysis of CO₂ Fixation. *Chemical Reviews*, **2013**. 113(8), 6621–6658.
4. Hori, Y. *Electrochemical CO₂ Reduction on Metal Electrodes*. In *Modern Aspects of Electrochemistry* (pp. 89–189). New York, NY: Springer New York 2008.
5. Nørskov, J. K., Bligaard, T., Rossmeisl, J., Christensen, C. H. Towards the computational design of solid catalysts. *Nature Chemistry*, **2009**. 1(1), 37–46.
6. Sundararaman, R., Goddard, W. A., Arias, T. A. Grand canonical electronic density-functional theory: Algorithms and applications to electrochemistry. *The Journal of Chemical Physics*, **2017**. 146(11), 114104.
7. Vesborg, P. C. K., Seger, B., Chorkendorff, I. Recent Development in Hydrogen Evolution Reaction Catalysts and Their Practical Implementation. *The Journal of Physical Chemistry Letters*, **2015**. 951–957.
8. Yang, J., Shin, H. S. Recent advances in layered transition metal dichalcogenides for hydrogen evolution reaction. *J. Mater. Chem. A*, **2014**. 2(17), 5979–5985.

9. Hinnemann, B., Moses, P. G., Bonde, J., Jørgensen, K. P., Nielsen, J. H., Horch, S., Nørskov, J. K. Biomimetic hydrogen evolution: MoS₂ nanoparticles as catalyst for hydrogen evolution. *Journal of the American Chemical Society*, **2005**. 127(15), 5308–5309.
10. Jaramillo, T., Jørgensen, K., Bonde, J. Identification of active edge sites for electrochemical H₂ evolution from MoS₂ nanocatalysts. *Science*, **2007**. 317(5834), 100–102.
11. Huang, Y., Nielsen, R. J., Goddard, W. A., Soriaga, M. P. The Reaction Mechanism with Free Energy Barriers for Electrochemical Dihydrogen Evolution on MoS₂. *J. Am. Chem. Soc.*, **2015**. 137 (20), pp 6692–6698.
12. Lukowski, M. A., Daniel, A. S., Meng, F., Forticaux, A., Li, L., Jin, S. Enhanced hydrogen evolution catalysis from chemically exfoliated metallic MoS₂ nanosheets. *Journal of the American Chemical Society*, **2013**. 135(28), 10274–10277.
13. Voiry, D., Salehi, M., Silva, R., Fujita, T., Chen, M., Asefa, T., Chhowalla, M. Conducting MoS₂ Nanosheets as Catalysts for Hydrogen Evolution Reaction. *Nano Letters*, **2013**. 13(12), 6222–6227.
14. Benck, J. D., Chen, Z., Kuritzky, L. Y., Forman, A. J., Jaramillo, T. F. Amorphous molybdenum sulfide catalysts for electrochemical hydrogen production: Insights into the origin of their catalytic activity. *ACS Catalysis*, **2012**. 2(9), 1916–1923.
15. Morales-Guio, C. G., Hu, X. Amorphous molybdenum sulfides as hydrogen evolution catalysts. *Accounts of Chemical Research*, **2014**. 47(8), 2671–2681.
16. Li, Y., Yu, Y., Huang, Y., Nielsen, R. A., Goddard, W. A., Li, Y., Cao, L. Engineering the Composition and Crystallinity of Molybdenum Sulfide for High-Performance Electrocatalytic Hydrogen Evolution. *ACS Catalysis*, **2015**. 5(1), 448–455.

17. Lassalle-Kaiser, B., Merki, D., Vrubel, H., Gul, S., Yachandra, V. K., Hu, X., Yano, J. Evidence from in situ X-ray absorption spectroscopy for the involvement of terminal disulfide in the reduction of protons by an amorphous molybdenum sulfide electrocatalyst. *Journal of the American Chemical Society*, **2015**. 137(1), 314–321.
18. Tan, Y., Liu, P., Chen, L., Cong, W., Ito, Y., Han, J., Chen, M. W. Monolayer MoS₂ films supported by 3D nanoporous metals for high-efficiency electrocatalytic hydrogen production. *Advanced Materials*, **2014**. 26(47), 8023–8028.
19. Li, H., Du, M., Mleczko, M. J., Koh, A. L., Nishi, Y., Pop, E., Zheng, X. Kinetic Study of Hydrogen Evolution Reaction over Strained MoS₂ with Sulfur Vacancies Using Scanning Electrochemical Microscopy. *Journal of the American Chemical Society*, **2016**. 138(15), 5123–5129.
20. Zhang, J., Wu, J., Guo, H., Chen, W., Yuan, J., Martinez, U., Lou, J. Unveiling Active Sites for the Hydrogen Evolution Reaction on Monolayer MoS₂. *Advanced Materials*, **2017**. 29(42), 1701955.
21. Gao, G., Sun, Q., Du, A. Activating catalytic inert basal plane of molybdenum disulfide to optimize hydrogen evolution activity via defect doping and strain engineering. *Journal of Physical Chemistry C*, **2016**. 120(30), 16761–16766.
22. Li, H., Tsai, C., Koh, A. L., Cai, L., Contryman, A. W., Fragapane, A. H., Zheng, X. Activating and optimizing MoS₂ basal planes for hydrogen evolution through the formation of strained sulphur vacancies. *Nature Materials*, **2015**. Volume 15, pages 48–53.
23. Li, G., Zhang, D., Qiao, Q., Yu, Y., Peterson, D., Zafar, A., Cao, L. All The Catalytic Active Sites of MoS₂ for Hydrogen Evolution. *Journal of the American Chemical Society*, **2016**. 138(51), 16632–16638.

24. Tsai, C., Li, H., Park, S., Park, J., Han, H. S., Nørskov, J. K., Abild-Pedersen, F. Electrochemical generation of sulfur vacancies in the basal plane of MoS₂ for hydrogen evolution. *Nature Communications*, **2017**. 8, 1–8.
25. Wiensch, J. D., John, J., Velazquez, J. M., Torelli, D. A., Pieterick, A. P., McDowell, M. T., Lewis, N. S. Comparative Study in Acidic and Alkaline Media of the Effects of pH and Crystallinity on the Hydrogen-Evolution Reaction on MoS₂ and MoSe₂. *ACS Energy Letters*, **2017**. 2234–2238.
26. Skúlason, E., Karlberg, G. S., Rossmeisl, J., Bligaard, T., Greeley, J., Jónsson, H., Nørskov, J. K. Density functional theory calculations for the hydrogen evolution reaction in an electrochemical double layer on the Pt(111) electrode. *Physical Chemistry Chemical Physics*, **2007**. 9(25), 3241–3250.
27. Taylor, C. D., Neurock, M. Theoretical insights into the structure and reactivity of the aqueous/metal interface. *Current Opinion in Solid State and Materials Science*, **2005**. 9(1–2), 49–65.
28. Taylor, C., Wasileski, S., Filhol, J.-S., Neurock, M. First principles reaction modeling of the electrochemical interface: Consideration and calculation of a tunable surface potential from atomic and electronic structure. *Physical Review B*, **2006**. 73(16), 165402.
29. Jinnouchi, R., Anderson, A. B. Electronic structure calculations of liquid-solid interfaces: Combination of density functional theory and modified Poisson-Boltzmann theory. *Physical Review B*, **2008**. 77(24), 245417.
30. Sha, Y., Yu, T. H., Merinov, B. V., Goddard, W. A. Prediction of the Dependence of the Fuel Cell Oxygen Reduction Reactions on Operating Voltage from DFT Calculations. *The Journal of Physical Chemistry C*, **2012**. 116(10), 6166–6173.
31. Gunceler, D., Letchworth-Weaver, K., Sundararaman, R., Schwarz, K. a, Arias, T. A. The importance of nonlinear fluid response in joint density-functional theory studies of

- battery systems. *Modelling and Simulation in Materials Science and Engineering*, **2013**. 21(7), 074005.
32. Sundararaman, R., Goddard, W. A. The charge-asymmetric nonlocally determined local-electric (CANDLE) solvation model. *The Journal of Chemical Physics*, **2015**. 142(6), 064107.
33. Goodpaster, J. D., Bell, A. T., Head-Gordon, M. Identification of Possible Pathways for C-C Bond Formation during Electrochemical Reduction of CO₂: New Theoretical Insights from an Improved Electrochemical Model. *Journal of Physical Chemistry Letters*, **2016**. 7(8), 1471–1477.
34. Xiao, H., Cheng, T., Goddard, W. A., Sundararaman, R. Mechanistic Explanation of the pH Dependence and Onset Potentials for Hydrocarbon Products from Electrochemical Reduction of CO on Cu (111). *J. Am. Chem. Soc.*, **2016**. 138 (2), pp 483–486.
35. Chan, K., Nørskov, J. K. Potential Dependence of Electrochemical Barriers from ab Initio Calculations. *Journal of Physical Chemistry Letters*, **2016**. 7(9), 1686–1690.
36. Janik, M. J., Taylor, C. D., Neurock, M. First-Principles Analysis of the Initial Electroreduction Steps of Oxygen over Pt(111). *Journal of The Electrochemical Society*, **2009**. 156(111), B126.
37. Morrison, S. R. *Electrochemistry at Semiconductor and Oxidized Metal Electrodes*, Springer U.S., Plenum Press, New York, 1980.
38. Bard, A. J., Faulkner, L. R. *Electrochemical Methods*, John Wiley & Sons, New York, 2001.
39. Sato, N. *Electrochemistry at Metal and Semiconductor Electrodes.*, Elsevier Science B. V., Amsterdam, The Netherlands, 1998.

40. v. Neumann, J. Zur Theorie der Gesellschaftsspiele. *Mathematische Annalen*, **1928**. 100(1), 295–320.
41. Zhou, H., Yu, F., Huang, Y., Sun, J., Zhu, Z., Nielsen, R. J., He, R., Bao, J., Goddard, W.A., Ren, Z. Efficient hydrogen evolution by ternary molybdenum sulfoselenide particles on self-standing porous nickel diselenide foam. *Nature Communications*, **2016**. 7, 12765.
42. Yu, F., Zhou, H., Huang, Y., Sun, J., Qin, F., Bao, J., Goddard, W.A., Chen, S., Ren, Z. High-performance bifunctional porous non-noble metal phosphide catalyst for overall water splitting. *Nature Communications*, **2018**. 9(1), 2551.

Supporting Information

1. Computational Method

In the GCP-K formulation, both constant charge or constant potential calculations can be used to obtain the parameters for the free energy $F(n)$ and GCP(U). In this study, we first calculated $F(n)$ at different charges, and then fitted for the quadratic form $F(n) = a(n-n_0)^2 + b(n-n_0) + c$, where a , b , and c are fitted parameters. Once these parameters are fitted, physical parameters are obtained by setting $c = F_0$, $b = \mu_{e,\text{SHE}} - eU_{\text{PZC}}$, and $a = -1/(2C_{\text{diff}})$. The last parameter, $\mu_{e,\text{SHE}}$, depending on the underlying quantum method, is described below. With these, the Equations 5(a) and 5(b) can be written out explicitly.

The free energies used in this study are calculated using the PBE-D3 functional [1-3], with geometry optimized in VASP [4-9], and a single point calculation at the end using jDFTx (v.1.2.1) [10] for the CANDLE solvation model [11].

The geometries are optimized in VASP, in which a conjugate gradient is used to search for the stable states, and the VTST package [12] is used to obtain the transition states. The VASPsol [13,14] solvation model is also used in VASP to include the effect of solvation. The wavefunction includes contributions up to a kinetic energy cutoff of 300eV. Since we use a 3x3 periodic slab of MoS₂, we use a gamma-center 4x4 Monkhorst-Pack k-point grid. The electronic energies are converged to 1E-5 eV, and the geometries are optimized to 0.01 eV/Å.

The single point joint density function functional theory calculation is used to obtain the combined DFT and solvation energy after the geometry is optimized. The CANDLE solvation model is used to describe solvation implicitly. Using this scheme, $\mu_{e,\text{SHE}} = 4.66$ eV, which is used for all structures. A kinetic energy cutoff of 13 Hartree is used for the wavefunction, and a k-point mesh of 4x4 is used. The free energies are converged to 1E-7 Hartree.

2. Parameters for the Grand Canonical Potential or Free Energy

Acid:

Species	a (Hartree/electron ²)	b (Hartree/electron)	c (Hartree)	vibrational contribution (ZPE + Hvib- Tsvib) Unit: kcal/mol
[MoS ₂] or 0	0.008566	-4.035820	- 323.556000	0.000000
[MoS ₂]H or 1	0.008773	-4.148660	- 308.916000	3.170408
[MoS ₂]H ₂ or 2	0.008606	-4.085280	- 314.974000	8.310925
[MoS ₂]H ₃ or 3	0.012046	-5.761100	- 111.340000	13.522096

Base:

Species	a (Hartree/electron ²)	b (Hartree/electron)	c (Hartree)	vibrational contribution (ZPE + Hvib- Tsvib) Unit: kcal/mol
[MoS ₂] or 0	0.008566	-4.035820	- 323.556000	0.000000
[MoS ₂]H or 1	0.008773	-4.148660	- 308.916000	3.170408
[MoS ₂]H ₂ or 2	0.008606	-4.085280	- 314.974000	8.310925
[MoS ₂]H ₃ or 3	0.012046	-5.761100	- 111.340000	13.522096

The complete kinetic model including all the parameters is attached at the end of the document. Python3 with the numpy library and the matplotlib library is needed to run the script.

3. References for the supporting information:

- [1] J. P. Perdew, K. Burke, and M. Ernzerhof. Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 77:3865, 1996.
- [2] J. P. Perdew, K. Burke, and M. Ernzerhof. Erratum: Generalized gradient approximation made simple. *Phys. Rev. Lett.*, 78:1396, 1997.
- [3] Grimme, S., Antony, J., Ehrlich, S., Krieg, H. (2010). A consistent and accurate ab initio parametrization of density functional dispersion correction (DFT-D) for the 94 elements H-Pu. *Journal of Chemical Physics*, 132(15).
- [4] G. Kresse and J. Hafner. Ab initio molecular dynamics for liquid metals. *Phys. Rev. B*, 47:558, 1993.
- [5] G. Kresse and J. Hafner. Ab initio molecular-dynamics simulation of the liquid-metal-amorphous-semiconductor transition in germanium. *Phys. Rev. B*, 49:14251, 1994.
- [6] G. Kresse and J. Furthmüller. Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set. *Comput. Mat. Sci.*, 6:15, 1996.
- [7] G. Kresse and J. Furthmüller. Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set. *Phys. Rev. B*, 54:11169, 1996.
- [8] P. E. Blochl. Projector augmented-wave method. *Phys. Rev. B*, 50:17953, 1994.
- [9] G. Kresse and D. Joubert. From ultrasoft pseudopotentials to the projector augmented-wave method. *Phys. Rev. B*, 59:1758, 1999.
- [10] R. Sundararaman, K. Letchworth-Weaver, K.A. Schwarz, D. Gunceler, Y. Ozhabes and T.A. Arias, 'JDFTx: software for joint density-functional theory', *SoftwareX* 6, 278 (2017)

- [11] Sundararaman, R., Goddard, W. A. (2015). The charge-asymmetric nonlocally determined local-electric (CANDLE) solvation model. *The Journal of Chemical Physics*, 142(6), 064107.
- [12] D. Sheppard, P. Xiao, W. Chemelewski, D. D. Johnson, and G. Henkelman, "A generalized solid-state nudged elastic band method," *J. Chem. Phys.* 136, 074103 (2012).
- [13] Mathew, K., Sundararaman, R., Letchworth-Weaver, K., Arias, T. A., Hennig, R. G. (2014). Implicit solvation model for density-functional study of nanocrystal surfaces and reaction pathways. *Journal of Chemical Physics*, 140(8).
- [14] Mathew, K., & Hennig, R. G. (2016). Implicit self-consistent description of electrolyte in plane-wave density-functional theory. *ArXiv*, 1–6.

IDENTIFICATION OF THE SELECTIVE SITES FOR ELECTROCHEMICAL REDUCTION OF CO TO C₂₊ PRODUCTS ON COPPER NANOPARTICLES BY COMBINING REACTIVE FORCE FIELDS, DENSITY FUNCTIONAL THEORY, AND MACHINE LEARNING

Abstract

Recent experiments have shown that CO reduction on oxide derived Cu nanoparticles (NP) is highly selective towards C₂₊ products. However, understanding of the active sites on such NPs is limited, because the NPs have ~200000 atoms with more than 10,000 surface sites, there are far too many for direct quantum mechanical calculations and experimental identifications. We show here how to overcome the computational limitation by combining multiple levels of theoretical computations with machine learning. This approach allows us to map the machine learned CO adsorption energies on the surface of the copper nanoparticle to construct the active site visualization (ASV). Furthermore, we identify the structural criteria for optimizing selective reduction by predicting the reaction energies of the potential determining step, ΔE_{OCCOH} , for C₂₊ product. Based on this structural criterion, we design a new periodic copper structure for CO reduction with a theoretical faradaic efficiency of 97%.

1. Introduction

Rapid progress is being made in developing new catalysts that are highly active and selective to electrochemically reduce CO or CO₂ to specific chemical fuels and feedstocks¹⁻². Improved selectivity and activity in reducing CO₂ and CO to valuable hydrocarbons and alcohols will enable the conversion of intermittent or remote renewable energies into complex chemical forms for storage and delivery³. At the same time, using sequestered CO₂ as the feedstock would reduce the amount of excess atmospheric CO₂ by completing the

carbon cycle with carbon fixation via artificial photosynthesis or other forms of renewable energy sources⁴⁻⁵.

However, CO₂ is quite stable, making it very challenging to optimize catalytic efficiency due to the difficulty in activating CO₂⁶. After decades of development, copper remains the only catalyst that can reduce CO or CO₂ by more than two electrons to generate valuable products in nontrivial amounts. Recently, oxide derived copper nanoparticles (NP) have been shown to greatly improve both the activity and selectivity of CO and CO₂ reduction towards C₂₊ products⁷. Based on early temperature programmed desorption (TPD) experiments, the improved performance of the oxide derived metal NP was hypothesized to arise from strong CO adsorption sites⁸. However, later experiments have found that selectivity correlated linearly with the grain boundary (GB) density⁹⁻¹⁰. In this work, we focus on elucidating which local Cu structures lead to the optimum properties for CO reduction to C₂₊ products.

2. Machine learning and CO adsorption

We previously used Density functional theory (DFT) with full solvent and Grand Canonical techniques to determine the reaction mechanisms for CO reduction to C1 and C2 products on Cu (100) and Cu (111) surfaces, leading to an excellent agreement with experiments (overpotentials within 0.05 V)¹¹⁻¹². However, the experimental 10nm NP involves ~200,000 atoms with ~10,000 possible surface sites, well beyond the capabilities of DFT. To circumvent the limitation of the direct application of DFT, we subsequently utilized the reactive force field (ReaxFF)¹³⁻¹⁴ to computationally grow the 10nm nanoparticles and then used DFT to sample only 84 surface sites for ΔE_{CO} and 4 surface sites for ΔE_{OCCOH} ¹⁵⁻¹⁶. In order to extract a quantitative understanding of the variations of the chemistry over the whole nanoparticle, we propose here a methodology to combine limited numbers of DFT calculations with machine learning to train a machine learning model that accurately predicts the binding energies for all sites.

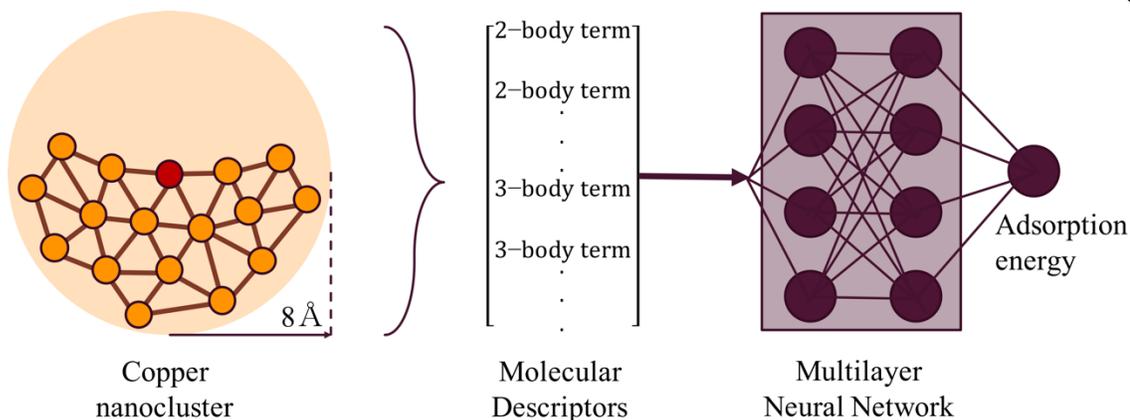


Figure 1. Schematics of the machine learning model. For each surface site (red), we extract a copper cluster including all the atoms within 8\AA from a copper NP. The adsorption energy is calculated using DFT and is used as the target property for training. We use Behler-Parrinello¹⁷ type neural network model. In this study, we describe the copper cluster by a set of 2-body and 3-body molecular descriptors about the surface atom. We then used these descriptors as input to a multilayer neural network for fitting.

First, we used ReaxFF to computationally synthesize a 10 nm copper nanoparticle (NP) that closely resembles the experimental NP [S1.1 in SI]. The predicted structure leads to XRD spectra and TEM images that match those of the experimental NP structures. Next, we selected 400 random surface sites and calculated their CO adsorption energies using DFT [S1.2 in SI]. We previously found that including atoms up to 8\AA from the surface site is sufficient to represent the local environment¹⁵. We integrated this local environment into a neural network in which the surrounding atoms are transformed into 12 two-body and 18 three-body molecular descriptors as inputs to a 2-layer neural network with 50 nodes in each layer, as shown in **Figure 1**. Further details of the descriptor definition are in Section S1.3 of the Supporting Information. We partitioned the 400 surface sites into training set, validation set, and test set with an 8:1:1 ratio. Here the validation set is used to terminate the training sufficiently early to avoid overfitting. Section S2 of the Supporting Information shows that the root mean squared error (RMSE) of the CO binding energy (ΔE_{CO}) on the training set is 0.111eV while for the validation set RMSE= 0.117eV, and for the test set RMSE= 0.123eV. We refer to this as the ReaxQM-Machine Learning strategy, or RxQM-

ML. This is much lower than the RMSE=0.2eV for a similar study of the crystalline surface of the NiGa binary alloy¹⁸.

After training this accurate neural network model, we used RxQM-ML to predict the CO adsorption energies for all 10,000 surface sites. The statistical distribution of the CO adsorption energies is shown in Figure 2(a). Overall, the CO adsorptions range from -0.55 to -1.43 eV, showing the wide variety of surface sites on the copper NP. As expected, most energies are clustered around the values for such low index surfaces, as (111), (100), and (211)¹⁵. However, we find a significant number of surface sites with much stronger CO adsorption energies. This is shown by the distribution to the left of the (211) line. These results are consistent with the TPD experiments, which show a broad peak centered at 275K only for the copper NP, indicating that ~7-15% of the surface leads to stronger CO adsorptions than low index copper surfaces⁸.

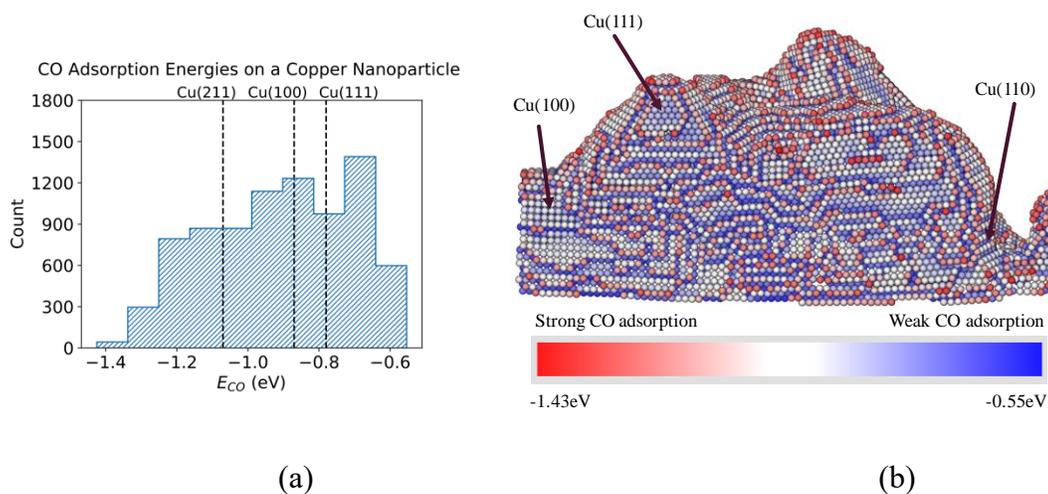


Figure 2. (a) Distribution of CO binding energies (ΔE_{CO}) on the 10nm copper nanoparticle. The three vertical dashed lines correspond to the CO adsorption energies of single crystal surfaces of (211), (100) and (111)¹⁵. (b) Active site visualization (ASV) of the predicted CO adsorption energies on the nanoparticle. As indicated by the colored bar, the red sites correspond to strong CO adsorption, the white sites correspond to moderate CO adsorption, and the blue sites correspond to weak CO adsorption. The common surfaces of (100), (111), and (110) are indicated in the figure.

Furthermore, the low-cost of RxQM-ML model makes it possible to establish the quantitative structure-activity relationship (QSAR) such that the machine-learned CO adsorption energies

can be remapped back to the copper nanoparticle, as shown in Figure 2(b). Here, red indicates low ΔE_{CO} , white indicates moderate ΔE_{CO} , and blue indicates unfavorable ΔE_{CO} . The (100), (111) and (110) surfaces are all colored light blue, indicating that they are near the mean values of the adsorption energy distributions as in Figure 2(a). The sites in solid blue are not fully exposed, making them difficult for CO to bind. The sites in red are of most interest because they correspond to more favorable adsorptions of CO than the low-index surfaces. As shown in light red in the figure, the moderately strong CO adsorption sites are typically along the step edges, and as shown in solid red, the strong CO adsorption sites are mostly isolated surface sites or kink sites.

The ASV in Figure 2 shows clearly that favorable CO adsorption sites are scattered across the whole nanoparticle surface. This is consistent with experimental observations that the surface areas corresponding to GBs are not sufficiently large to account for the number of strong CO adsorption sites⁸. Using RxQM-ML, we now directly show that the strong CO adsorption energies are not just at GBs.

3. C2 coupling

Although we have demonstrated that the CO binding energy is not necessary to correlate with GBs, there is a great deal of experimental evidence suggesting that increasing the GB density can significantly improve the C₂₊ selectivity. Another descriptor is needed to describe selectivity of these nanoparticles. As shown experimentally¹⁹ and theoretically²⁰⁻²², the selective step towards C₂₊ products involves C-C coupling in which *OCCOH is formed. Thus, the most plausible descriptor is the reaction energy for forming *OCCOH,

$$\Delta E_{\text{OCCOH}} = E[*\text{OCCOH}] - E[*\text{CO}, *\text{CO}] - 0.5 \times E[\text{H}_2],$$

which we have shown previously to be the potential determining step for ethanol production.

Then, we started with ~180 randomly sampled surface site and calculated the formation energy for *OCCOH, ΔE_{OCCOH} . The distribution is shown in the blue histograms in Figure 3(a). As shown in the figure, the range of ΔE_{OCCOH} spans by more than 1eV, implying that

some sites are much more selective than others. We could sample additional sites to develop a similar machine learning model for ΔE_{OCCOH} . However, we examined the sites with the lowest ΔE_{OCCOH} , and found that all of them are involved square sites, similar to those of the (100) surface. To test this hypothesis, we further sampled 100 *square* sites, leading to the distribution for ΔE_{OCCOH} shown in orange in **Figure 3(a)**. Comparing to the random sites on the surface of the copper nanoparticle, we found that the square sites were indeed more favorable, as shown by the shift in the distribution in ΔE_{OCCOH} in **Figure 3(a)**.

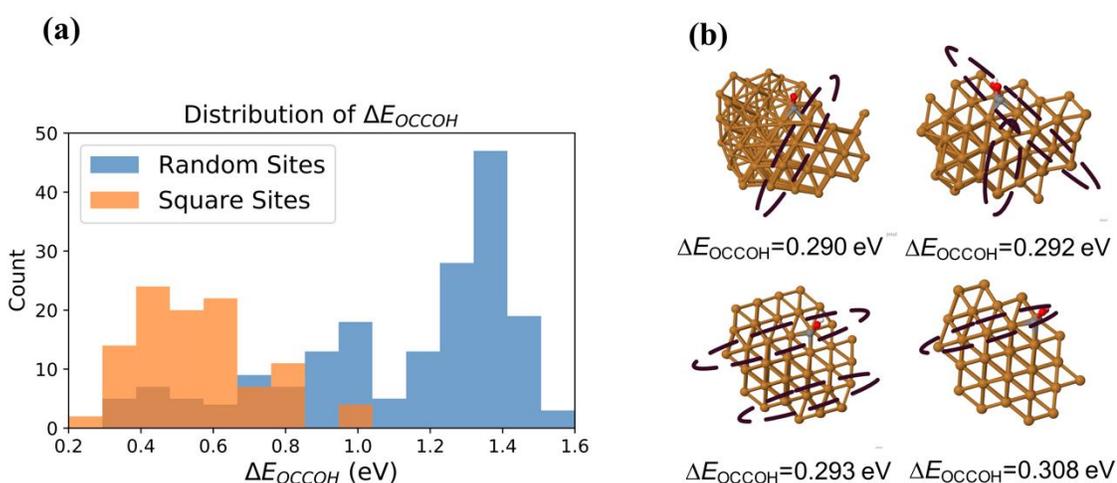


Figure 3. (a) Distributions of ΔE_{OCCOH} on the surface of the copper nanoparticle. Blue: 180 *random* surface sites; orange: 100 random *square* sites. (b) The four square structures with the lowest ΔE_{OCCOH} sampled randomly from the copper nanoparticle. The dashed ellipses indicate the locations of the twin boundaries.

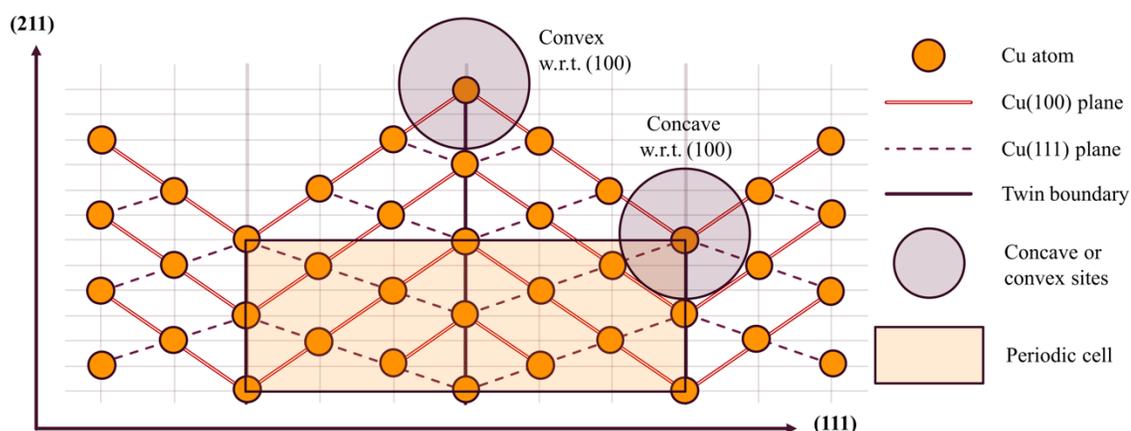
With the new distribution of just the *square* sites, we extracted the common features of the most selective sites by further examining the square sites with the lowest ΔE_{OCCOH} . We found that a step (111) surface is always next to these favorable square sites, as shown in Figure 3(b). These sites are similar to the Cu(S)[$n(100) \times (111)$] edge step sites where the (111) surface and the (100) surface intersect. In fact, experiments²³⁻²⁴ showed that these step sites have higher selectivity than either the (100) and (111) surface. To confirm this theoretically, we calculated E_{OCCOH} on (100), (111), (311), and (511) surfaces to be 0.44eV, 0.64eV, 0.52eV and 0.41eV. The calculated trend agrees very well the experimental selectivity trend

in which $(511) > (100) > (311) > (111)$ [Section **S3** in SI]. This comparison with experimental findings on the $\text{Cu(S)}[n(100) \times (111)]$ surfaces confirms the validity of using E_{OCCOH} as the descriptor for the selectivity towards C_{2+} products. It is also consistent with our finding from sampling the NP that favorable sites for E_{OCCOH} or C_{2+} selectivity must involve a (111) step surface next to a (100) site where $^*\text{OCCOH}$ is formed.

In addition, twin boundaries are associated with the square surface sites having the lowest ΔE_{OCCOH} . **Figure 3(b)** shows that these twin boundaries are all next to the site where $^*\text{OCCOH}$ is formed. This implies that the selectivity towards C_{2+} products is directly related to twin boundaries which are a special type of GBs.

Building on the idea that the above common features lead to the best $^*\text{OCCOH}$ sites, we constructed the smallest periodic structure possessing these features. This is shown in **Figure 4**. We expect that this periodic surface will behave chemically in the same way as these selective sites. Because it is a smallest periodic structure containing these sites, the density of active sites will be much higher than the randomly and sparsely distributed active sites on a nanoparticle.

(a):



(b):

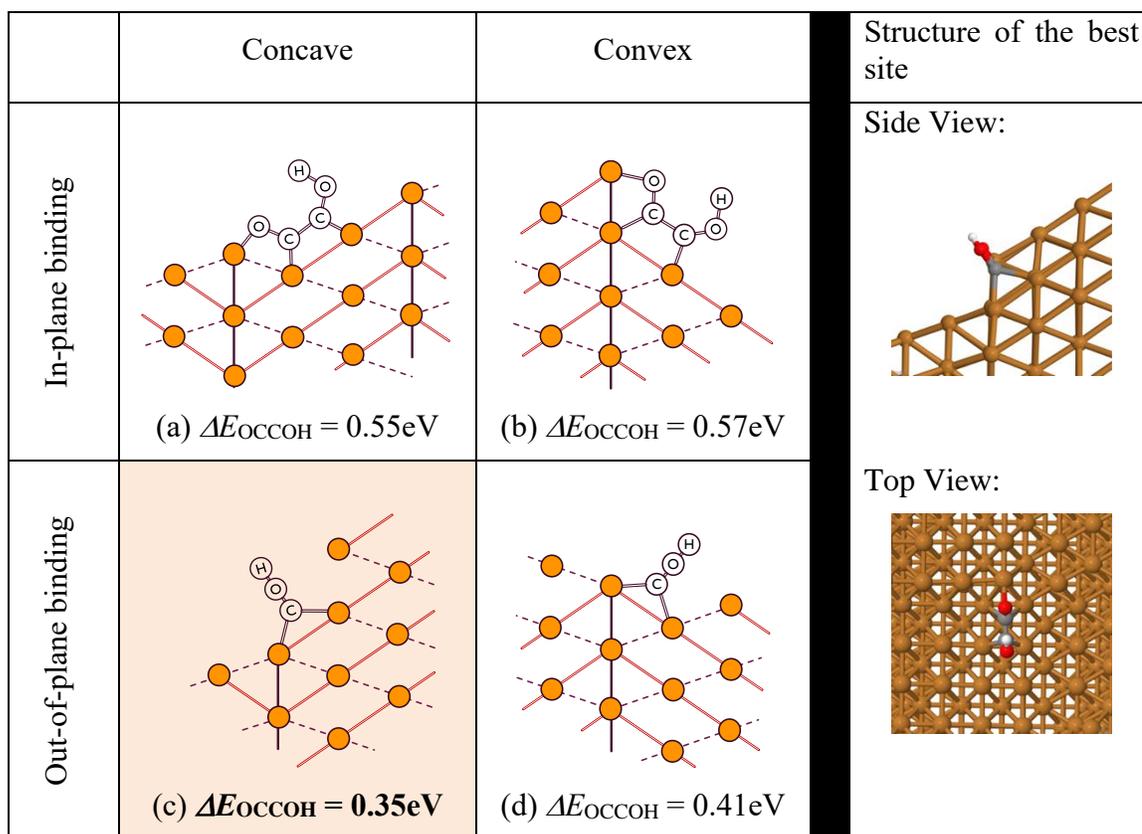


Figure 4. (a) The shaded area is the minimal periodic structure of FCC copper containing the (100) planes, (111) planes, and twin boundaries. Terminating this structure to expose the (100) and (111) surfaces leads to sites that are concave or convex with respect to the (100) planes. (b) The four types of sites for adsorbed $^*\text{OCCOH}$ on the surface of the minimal periodic structure. The structure that is concave with respect to the (100) planes has the most favorable ΔE_{OCCOH} for C_{2+} selectivity. The top and side views of this structure are shown on the right column of (b). More details on these sites are shown in Section S4 of the supporting information. Note, for structures (c) and (d), the other carbon and oxygen atoms are not shown since they overlap with the foreground atoms in the side view. The full $^*\text{OCCOH}$ structure for (c) is revealed in the top view, as shown on the right column.

This minimal periodic structure is shown in Figure 4(a). From the ABC stacking of the FCC copper, the smallest grain size must contain at least 6-layer, corresponding to ABCACB

stacking, where the A layers are twin boundaries. Since the step surfaces involving the (100) and (111) are of interest, they are shown by double lines and dash lines in the figure.

Based on the configurations of the adsorbed *OCCOH on the copper clusters, there are 4 ways of placing the intermediate on this surface, as shown in Figure 4(b). The first two structures, (a) and (b), with the *OCCOH adsorbed in the cross sectional plane show unfavorable energies. Thus, the in-plane *OCCOH adsorption is not responsible for the increased in C_{2+} selectivity. On the other hand, *OCCOH adsorbed perpendicular to the page (or out-of-plane) are much more favorable, with only 0.41 eV for the convex site and 0.35 eV for the concave site, which is better than all the single crystal surface sites considered here. In fact, the same configuration is also found for the copper nanoparticle. As shown again in Figure 3(b), *OCCOH are all adsorbed perpendicular to the page. Thus, we predict that the (100)-like square sites next to a (111)-like step surface and on-top of a twin boundary that binds *OCCOH parallel to the twin boundary will have the most favorable E_{OCCOH} , which corresponds to the most selective sites.

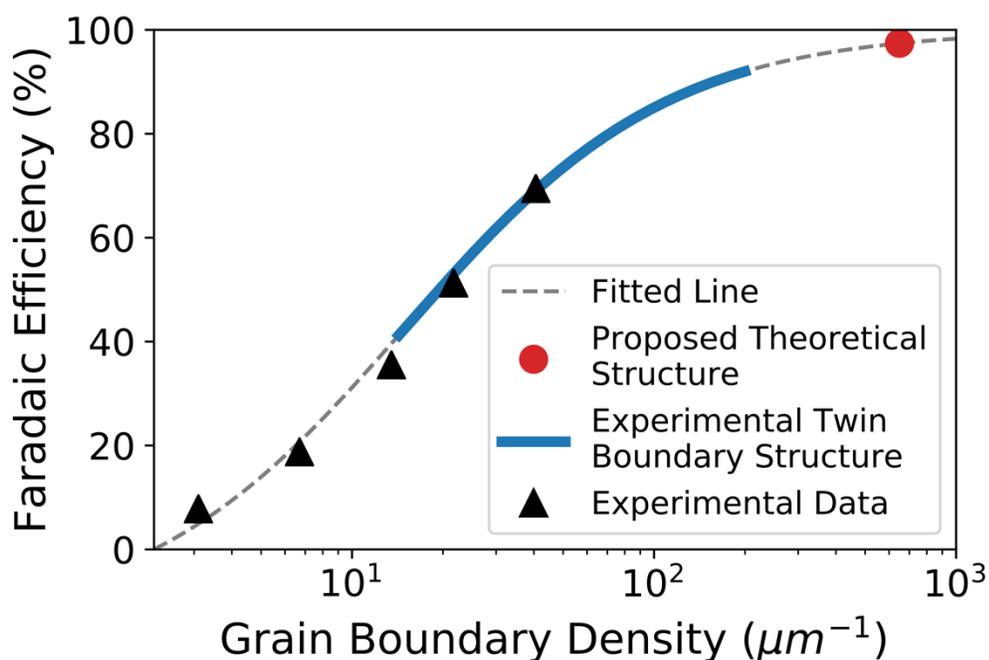


Figure 5. Predicted faradaic efficiency (FE) of the concave site on the minimal periodic structure compared to experimental data. An experimental copper structure with abundant twin boundaries²⁵ is also extrapolated based on the densities of the boundaries.

For the most favorable structure, the faradaic efficiency towards C₂₊ product is predicted to be 97% using experimental data in which the current density for C₂₊ production increases linearly as the density of GBs, and the current for hydrogen production remains the same¹⁰. The prediction is shown in **Figure 5**, which also includes a prediction for an experimental copper structure in which a high density of twin boundaries is synthesized²⁵, assuming that the structure exposes the twin boundaries in the preferable configuration. Details of this prediction are summarized in section **S4** of the Supporting Information.

4. Conclusion

In conclusion, we used machine learning to fit the structure-activity relationship between the local structures of the copper nanoparticle and the theoretical CO adsorption energies. By extrapolating the energies back to the nanoparticle, we found that strong CO adsorption energies are not just on GBs, implying that CO adsorption energies are not an appropriate descriptor for C₂₊ selectivity. Rather, we show that ΔE_{OCCOH} , the transition state for forming ethanol of C₂₊ products in C-C coupling, is the appropriate descriptor. This explains the selectivity on Cu(S)[n(100) × (111)] surfaces and the twin-related step square sites on the nanoparticle. To illustrate how to use this information, we designed the minimal periodic structure. This minimal periodic structure has a super high density of selective sites that we expect will lead to near unity selectivity based on extrapolations of theoretical and experimental data.

Reference:

1. Recent Zhang, W.; Hu, Y.; Ma, L.; Zhu, G.; Wang, Y.; Xue, X.; Chen, R.; Yang, S.; Jin, Z. Progress and Perspective of Electrocatalytic CO₂ Reduction for Renewable Carbonaceous Fuels and Chemicals. *Adv. Sci.* **2018**, 5(1), 1700275.

2. Zhao, G.; Huang, X.; Wang, X.; Wang, X. Progress in Catalyst Exploration for Heterogeneous CO₂ Reduction and Utilization: A Critical Review. *J. Mater. Chem. A*. **2017**, 5(41), 21625–21649.
3. Detz, R. J.; Reek, J. N. H.; Van Der Zwaan, B. C. C. The Future of Solar Fuels: When Could They Become Competitive? *Energy Environ. Sci.* **2018**, 11(7), 1653–1669.
4. Van Der Giesen, C.; Kleijn, R.; Kramer, G. J. Energy and Climate Impacts of Producing Synthetic Hydrocarbon Fuels from CO₂. *Environ. Sci. Technol.* **2014**, 48(12), 7111–7121.
5. Artz, J.; Müller, T. E.; Thenert, K.; Kleinekorte, J.; Meys, R.; Sternberg, A.; Bardow, A.; Leitner, W. Sustainable Conversion of Carbon Dioxide: An Integrated Review of Catalysis and Life Cycle Assessment. *Chem. Rev.* **2018**, 118(2), 434–504.
6. Kumar, B.; Llorente, M.; Froehlich, J.; Dang, T.; Sathrum, A.; Kubiak, C. P. Photochemical and Photoelectrochemical Reduction of CO₂. *Annu. Rev. Phys. Chem.* **2012**, 63, 541–569.
7. Li, C. W.; Kanan, M. W. CO₂ Reduction at low Overpotential on Cu Electrodes Resulting from the Reduction of Thick Cu₂O Films. *J. Am. Chem. Soc.* **2015**, 134(17), 7231–7234.
8. Verdaguier-Casadevall, A.; Li, C. W.; Johansson, T. P.; Scott, S. B.; McKeown, J. T.; Kumar, M.; Stephens, I. E. L.; Kanan, M. W.; Chorkendorff, I. Probing the Active Surface Sites for CO Reduction on Oxide-Derived Copper Electrocatalysts. *J. Am. Chem. Soc.* **2015**, 137(31), 9808–9811.
9. Feng, X.; Jiang, K.; Fan, S.; Kanan, M. W. Grain-Boundary-Dependent CO₂ Electroreduction Activity. *J. Am. Chem. Soc.* **2015**, 137(14), 4606–4609.

10. Feng, X.; Jiang, K.; Fan, S.; Kanan, M. W. A Direct Grain-Boundary-Activity Correlation for CO Electroreduction on Cu Nanoparticles. *ACS Cent. Sci.* **2016**, *2*(3), 169–174.
11. Nørskov, J. K.; Bligaard, T.; Rossmeisl, J.; Christensen, C. H. (2009). Towards the Computational Design of Solid Catalysts. *Nature Chemistry*. **2009**, *1*(1), 37–46.
12. Becke, A. D. Perspective: Fifty Years of Density-Functional Theory in Chemical Physics. *J. Che. Phys.* **2014**, *140*(18), 18A301.
13. van Duin, A. C. T.; Dasgupta, S.; Lorant, F.; Goddard, W. A. (2001). ReaxFF: A Reactive Force Field for Hydrocarbons. *J. Phys. Chem. A.* **2001**, *105*(41),
14. Chenoweth, K.; van Duin, A. C. T.; Goddard, W. A. ReaxFF Reactive Force Field for Molecular Dynamics Simulations of Hydrocarbon Oxidation. *J. Phys. Chem. A.* **2008**, *112*(5), 1040–1053.
15. Cheng, T.; Xiao, H.; Goddard, W. A. Nature of the Active Sites for CO Reduction on Copper Nanoparticles; Suggestions for Optimizing Performance. *J. Am. Chem. Soc.* **2017**, *139*(34), 11642–11645.
16. Cheng, T.; Huang, Y.; Xiao, H.; Goddard, W. A. Predicted Structures of the Active Sites Responsible for the Improved Reduction of Carbon Dioxide by Gold Nanoparticles. *J. Phys. Chem. Lett.* **2017**, *8*(14), 3317–3320.
17. Behler, J.; Parrinello, M. Generalized Neural-Network Representation of High-Dimensional Potential-Energy Surfaces. *Phys. Rev. Lett.* **2007**, *98*(14), 1–4.
18. Ulissi, Z. W.; Tang, M. T.; Xiao, J.; Liu, X., Torelli, D. A.; Karamad, M.; Cummins, K.; Hahn, C.; Lewis, N. S.; Jaramillo, T. F.; Chan, K.; Nørskov, J. K. Machine-Learning Methods Enable Exhaustive Searches for Active Bimetallic Facets and Reveal Active Site Motifs for CO₂ Reduction. *ACS Catal.* **2017**, *7*(10), 6600–6608.

19. Pérez-Gallent, E.; Figueiredo, M. C.; Calle-Vallejo, F.; Koper, M. T. M. Spectroscopic Observation of a Hydrogenated CO Dimer Intermediate During CO Reduction on Cu(100) Electrodes. *Angew. Chem. Int. Ed.* **2017**, *56*(13), 3621–3624.
20. Calle-Vallejo, F.; Koper, M. T. M. Theoretical Considerations on the Electroreduction of CO to C₂ Species on Cu(100) Electrodes. *Angew. Chem. Int. Ed.* **2013**, *125*(28), 7423–7426.
21. Cheng, T.; Xiao, H.; Goddard, W. A. Full Atomistic Reaction Mechanism with Kinetics for CO Reduction on Cu(100) from ab initio Molecular Dynamics Free-Energy Calculations at 298 K. *Proc. Natl. Acad. Sci.* **2017**, *114*(8), 1795–1800.
22. Xiao, H.; Cheng, T.; Goddard, W. A. Atomistic Mechanisms Underlying Selectivities in C₁ and C₂ Products from Electrochemical Reduction of CO on Cu(111). *J. Am. Chem. Soc.* **2017**, *139*(1), 130–136.
23. Hori, Y.; Takahashi, I.; Koga, O.; Hoshi, N. Selective Formation of C₂ Compounds from Electrochemical Reduction of CO₂ at A Series of Copper Single Crystal Electrodes. *J. Phys. Chem. B.* **2002**, *106*(1), 15–17.
24. Kim, Y. G.; Javier, A.; Baricuatro, J. H.; Torelli, D.; Cummins, K. D.; Tsang C. F.; Hemminger, J. C.; Soriaga, M. P. Surface Reconstruction of Pure-Cu Single-Crystal Electrodes under CO-Reduction Potentials in Alkaline Solutions: A Study by Seriatim ECSTM-DEMS. *Journal of Electroanalytical Chemistry.* **2017**, *793*, 113–118.
25. Morris Wang, Y.; Sansoz, F.; Lagrange, T.; Ott, R. T.; Marian, J.; Barbee, T. W.; Hamza, A. V. Defective Twin Boundaries in Nanotwinned Metals. *Nature Materials.* **2013**, *12*(8), 697–702.

Supporting Information

References in the Supporting Information:

- [1] Cheng, T., Xiao, H., and Goddard, W. A. (2017). *Nature of the Active Sites for CO Reduction on Copper Nanoparticles; Suggestions for Optimizing Performance*. Journal of the American Chemical Society, 139(34), 11642–11645.
- [2] Daw, M. S., Foiles, S. M., Baskes, M. I. (1993). *The embedded-atom method: a review of theory and applications*. Materials Science Reports, 9(7–8), 251–310.
- [3] van Duin, A. C. T., Dasgupta, S., Lorant, F., and Goddard, W. A. (2001). *ReaxFF: A Reactive Force Field for Hydrocarbons*. The Journal of Physical Chemistry A, 105(41), 9396-9409.
- [4] G. Kresse and J. Hafner. *Ab initio molecular-dynamics simulation of the liquid-metal-amorphous-semiconductor transition in germanium*. Phys. Rev. B, 49:14251, 1994; G. Kresse and J. Furthmüller. *Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave basis set*. Comput. Mat. Sci., 6:15, 1996.; G. Kresse and J. Furthmüller. *Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set*. Phys. Rev. B, 54:11169, 1996.;
- [5] J. P. Perdew, K. Burke, and M. Ernzerhof. *Generalized gradient approximation made simple*. Phys. Rev. Lett., 77:3865, 1996; J. P. Perdew, K. Burke, and M. Ernzerhof. *Erratum: Generalized gradient approximation made simple*. Phys. Rev. Lett., 78:1396, 1997.
- [6] Behler, J., and Parrinello, M. (2007). *Generalized neural-network representation of high-dimensional potential-energy surfaces*. Physical Review Letters, 98(14), 1–4.

- [7] Cheng, T., Huang, Y., Xiao, H., and Goddard, W. A. (2017). *Predicted Structures of the Active Sites Responsible for the Improved Reduction of Carbon Dioxide by Gold Nanoparticles*. *Journal of Physical Chemistry Letters*, 8(14), 3317–3320.
- [8] Glorot, X. and Bengio, Y. (2010). *Understanding the difficulty of training deep feedforward neural networks*. *Proceedings of the 13th International Conference on Artificial Intelligence and Statistics*, 9, 249–256.
- [9] Hori, Y., Takahashi, I., Koga, O., and Hoshi, N. (2002). *Selective formation of C₂ compounds from electrochemical reduction of CO₂ at a series of copper single crystal electrodes*. *Journal of Physical Chemistry B*, 106(1), 15–17.
- [10] Feng, X., Jiang, K., Fan, S., and Kanan, M. W. (2016). *A Direct Grain-Boundary-Activity Correlation for CO Electroreduction on Cu Nanoparticles*. *ACS Central Science*, 2(3), 169–174.
- [11] Morris Wang, Y., Sansoz, F., Lagrange, T., Ott, R. T., Marian, J., Barbee, T. W., Hamza, A. V. (2013). *Defective twin boundaries in nanotwinned metals*. *Nature Materials*, 12(8), 697–702.

S1. Methods

S1.1 ReaxFF

The copper nanoparticle investigated in this study is computationally synthesized in the same way as in [1]. The embedded-atom model (EAM) [2] was used to model the interactions between copper atoms. To simulate the chemical vapor deposition (CVD) synthesis, copper atoms are added to the simulation box in a rate of 3.2A/ns for 30ns. After the nanoparticle condenses, a total of 38 simulated annealing cycles are applied to relax the initially formed structure. Each cycle involves a heating step from 300K to 1200K in 5ps, high temperature step at 1200K for 5ps, cooling step from 1200K to 300K in 5ps, and room

temperature step at 300K. Last, the final copper nanoparticle is relaxed using the reactive force field [3] at 300K for 20ps.

SI.2 DFT

The density functional theory (DFT) calculations are performed in VASP [4]. The PBE functional [5] is used to obtain electronic energy. The plane wave basis up to a kinetic energy cutoff of 400eV is used to describe the wave function. Since copper is metallic, the method of Methfessel-Paxton of order 1 is used for smearing. The smearing width is 0.2eV. The convergence criteria for electronic energies is 1e-5eV, and the geometries are optimized until the force converged to be within an error of 0.01eV/Å.

ΔE_{CO} and ΔE_{OCCOH} are calculated in the same way as in [1]:

$$\Delta E_{CO} = E[*CO] - E[*] - E[CO]$$

$$\Delta E_{OCCOH} = E[*OCCOH] - E[*CO, *CO] - 0.5 \times E[H_2]$$

SI.3 Neural Network Machine Learning Model

The neural network used in this study is of the type by Behler and Parrinello [6]. However, instead of summing up contributions from all atoms, only the contribution from the target surface atom is used. In this study, instead of the atomic energy, the quantity to be fitted using machine learning is the CO adsorption energy.

As we have shown earlier [1, 7], the CO adsorption energy can be sufficiently described on the surface of the nanoparticle by considering all the atoms within 8Å of the surface site. Thus, a copper cluster of 8Å radius about each surface atom is extracted from the nanoparticle to compute the CO adsorption energy. Subsequently, the same cluster is used as the input to the machine learning model. Since we are only considering the chemical behavior of the chosen surface site, only molecular descriptions about that site are used as input to the neural network. Similar to our previous model, two-body and three-body

features are used. The overall structure of the neural network machine learning model is shown as below:

And mathematically, the above model is represented as follows:

$$G_{i\alpha}^{(2)} = \sum_j \phi_{\alpha}^{(2)}(\mathbf{p}_{ij}), G_{i\beta}^{(3)} = \sum_{jk} \phi_{\beta}^{(3)}(\mathbf{p}_{ij}, \mathbf{q}_{ijk}) \quad (1)$$

$$E_{CO} = F_{NN}(\{G_{i\alpha}^{(2)}, G_{i\beta}^{(3)}\}; \mathbf{w}, \mathbf{b}) \quad (2)$$

Where $G_{i\alpha}^{(2)}$ and $G_{i\beta}^{(3)}$ are the 2- and 3-body descriptors from the 2- and 3-body geometrical features \mathbf{p}_{ij} and \mathbf{q}_{ijk} , e.g., the 2-body term can include the interatomic distance $p_{ij} = R_{ij}$, and the 3-body term can include the angle formed by three atoms $q_{ijk} = \theta_{ijk}$. Finally, the function $F_{NN}(\mathbf{G}; \mathbf{w}, \mathbf{b})$ represents the network that transforms the input descriptors \mathbf{G} into a single value using the parameters with weights \mathbf{w} and biases \mathbf{b} . The final value calculated from the whole model is the CO adsorption energy, as represented by equation (2).

The molecular descriptors as input to the neural network can be extracted in many ways. In this study, the set of piecewise cosine functions is used because it is a more systematic way of constructing locally based symmetry functions.

$$R_{\alpha}^k = R_{inner} + (\alpha - 1)h_k \quad \text{where } \alpha = 1, 2, \dots, M_k$$

$$\varphi_{\alpha}^{(k)}(R_{ml}) = \begin{cases} \frac{1}{2} \cos\left(\frac{R_{ml} - R_{\alpha}^k}{h_k} \pi\right) + \frac{1}{2}, & |R_{ml} - R_{\alpha}^k| < h_k \\ 0, & \text{Otherwise} \end{cases} \quad (3)$$

$$G_{\alpha,l}^{(2)} = \sum_m \varphi_{\alpha}^{(2)}(R_{ml}) \quad (4)$$

$$G_{\alpha\beta\gamma,l}^{(3)} = \sum_{m,n} \varphi_{\alpha}^{(3)}(R_{ml}) \varphi_{\beta}^{(3)}(R_{nl}) \varphi_{\gamma}^{(3)}(R_{mn}) \quad (5)$$

where:

$\varphi_{\alpha}^{(k)}(R_{ml})$ is the piecewise cosine function over the interatomic distance R_{ml}

$k=2,3$ corresponds the 2-body and 3-body terms, and

$h_k = (R_{\text{outer}} - R_{\text{inner}})/M_k$ is the width of the piecewise cosine functions.

The set of piecewise cosine functions are described by 4 quantities: the inner cutoff R_{inner} , the outer cutoff R_{outer} , the number of two-body functions M_2 , and the number of three-body functions M_3 . Here, the $M_2=12$ and $M_3=3$ are used, corresponding to a set of 30 unique descriptors for each copper cluster. Because each cluster is only of 8 angstroms, we can include all the atoms in the descriptor. Thus, R_{outer} is set to 8Å. Since only 12 two-body functions are used, the width of each piecewise cosine function is relatively large, thus, R_{inner} can be set to 0. Once the input atomic descriptors are assembled about the target surface site, it is fed through a two-layer neural network each with 50 nodes to produce one output, which is then fitted to the adsorption energy.

S2. Implementation and Training of the Neural Network Model

The above neural network is implemented in Python and is freely available on our Gitlab page at: https://gitlab.com/yufeng.huang/cunp_coads. Instructions on how to use the neural network model is described on the website.

Once the structure of the neural network is constructed, the weights and biases of the neural network model are initialized using the Xavier initializer [8]. Then using the 8:1:1 split of the data set for training, validation and testing, we obtain the following RMSE as a function of training iterations, as shown in Figure S2(a). The final RMSE of the training set, validation set, and test set are 0.111eV, 0.117eV and 0.123eV. The overall distributions between the DFT energies and the neural network predicted values also agree, as shown in Figure S2(b).

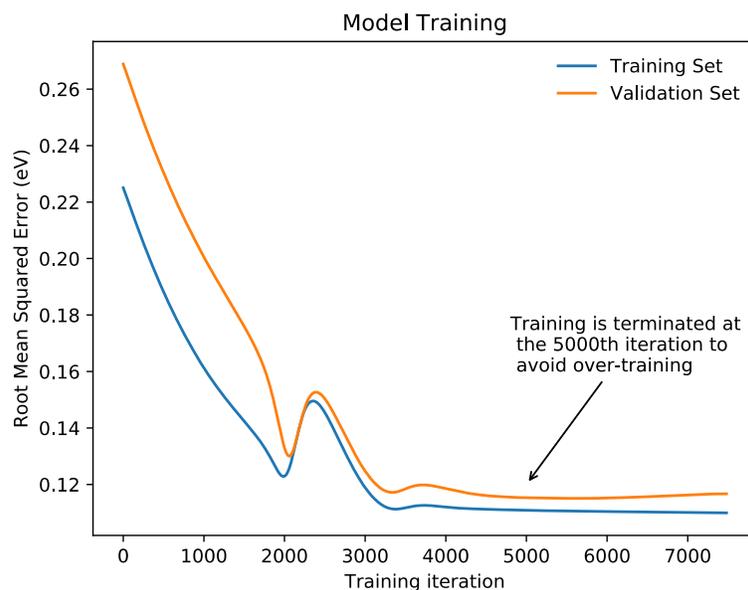


Figure S1. RMSE of the training set and validation set as functions of the training iterations.

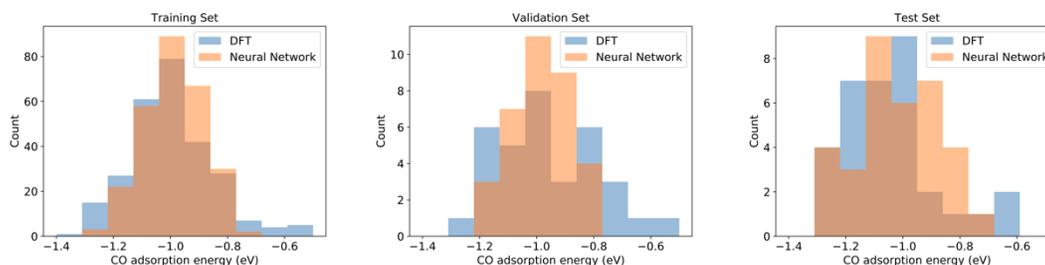


Figure S2. Distribution of the energies between the DFT values and the neural network predicted values for the training set, validation set, and the test set.

Unit: eV	DFT range	NN range
Training Set	[-1.284, -0.458]	[-1.262, -0.715]
Validation Set	[-1.276, -0.543]	[-1.158, -0.820]
Test Set	[-1.314, -0.433]	[-1.278, -0.761]

Table S1. Energy ranges between DFT and neural network for the training set, validation set, and test set. Because CO does not adsorb on certain unfavorable sites, these sites are treated as high adsorption energies. Due to this, the upper end of the range has much larger error to include the unfavorable outliers than the lower end. However, NN predicts the lower end of the ranges very well, indicating that good CO adsorption sites are well identified.

NN sizes	Training set RMSE	Validation set RMSE	Test set RMSE
30x2	0.110 eV	0.123	0.140
40x2	0.115	0.118	0.135
50x2	0.108	0.119	0.122
60x2	0.103	0.116	0.117

Table S2. RMSEs of the training set, validation set, and test set as functions of the neural network sizes. The notation $nx2$ indicates two hidden layers with n nodes are used in the neural network. Although the training set and validation set are fitted in similar accuracies between different sizes, the test set error is smaller when the size of the neural network increases. The neural network size of 50x2 is used to allow enough flexibility.

S3. Relationship between E_{OCCOH} and C_2 products selectivity

Hori et al. (2002) [9] reported the selectivity for C_2 products on different crystalline surfaces of copper. Using their reported values and our theoretical *OCCOH reaction energy, E_{OCCOH} , we found that the log of the ratio between C_2H_4 and CH_4 correlates almost linearly E_{OCCOH} , as shown in the following figure. This is not surprising because the reaction rates are typically exponentially dependent on energy as in the Eyring equation or Arrhenius equation. However, the good fit between the selectivity and E_{OCCOH} implies that E_{OCCOH} is a good descriptor for the selectivity of C_2 products.

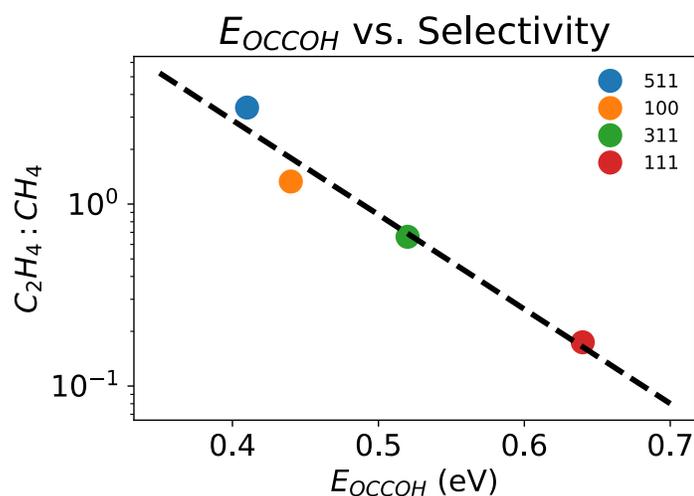


Figure S3. Experimental $\text{C}_2\text{H}_4:\text{CH}_4$ ratio and the reaction energy of E_{OCCOH} . Note that the logarithm of the ratios is used for the y-axis.

S4. Surface terminations of the Twin boundaries

Based on the configurations of the adsorbed *OCCOH on the copper clusters, there are 4 ways of placing the intermediate on the surface, as shown in the following figure, Figure S4. This is due to two factors. First, the (100) surfaces intersecting at the twin boundary can be concave or convex, as indicated by the purple figure in Figure 4(a) in the main text and the side view in Figure S4. Second, the plane formed by the adsorbed *OCCOH can be in the same plane as the page, or perpendicular to the page, as shown in top view in Figure S4.

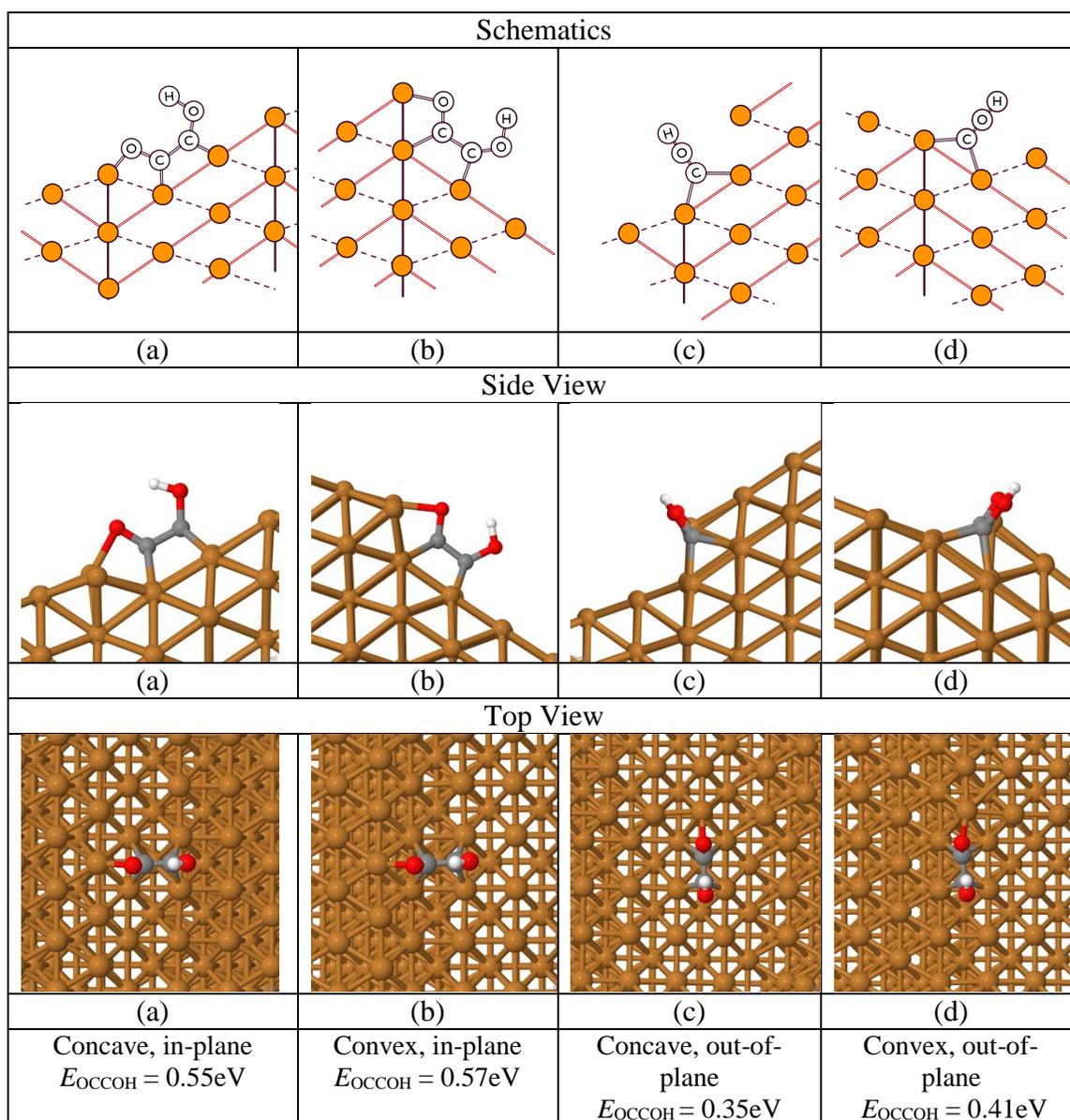


Figure S4. The four configurations of *OCCOH binding at the twin boundary of copper. (a): *OCCOH adsorbed in-plane at the concave site; (b): *OCCOH adsorbed in-plane at the convex site; (c) *OCCOH adsorbed out-of-plane at the concave site; (d) *OCCOH adsorbed out-of-plane at the convex site.

S5. Prediction of Faradaic Efficiencies

The data for currents $J_{C_{2+}}$ and J_{H_2} as functions of the grain boundary density d_{GB} are obtained from Figure 3(d) and (g) of reference [10]. First, since J_{H_2} does not depend on d_{GB} , and the terminal value is 3.1 mA/cm² for large d_{GB} , we then keep J_{H_2} constant for this approximation. For C_{2+} production, $J_{C_{2+}}$ is fitted linearly as in the experimental data. Then we have:

$$J_{C_{2+}} = 0.18 * d_{GB} - 0.399$$

$$J_{H_2} = 0.31$$

Where d_{GB} is in units of μm^{-1} , $J_{C_{2+}}$ and J_{H_2} are in units of mA/cm².

Our periodic structure has twin boundary density of 649.5 μm^{-1} . Thus, the faradaic efficiency is then $J_{C_{2+}} / (J_{C_{2+}} + J_{H_2}) = 97.4\%$

Also, [11] reported a copper structure with a large abundance of twin boundary sites with spacing of 5~70nm, which corresponds to densities of 14.3 to 200 μm^{-1} . The corresponding faradaic efficiencies are 41.2% to 92.0%.

The results are summarized in Figure 5 of the main text.

DFT BASED NEURAL NETWORK FORCE FIELDS FROM ENERGY DECOMPOSITIONS

Abstract

In order to develop force fields (FF) for molecular dynamics simulations that retain the accuracy of *ab initio* density functional theory (DFT), we developed a machine learning (ML) protocol based on an energy decomposition scheme that extracts atomic energies from DFT calculations. Our DFT to FF (DFT2FF) approach provides almost hundreds of times more data for the DFT energies, which dramatically improves accuracy with less DFT calculations. In addition, we use piecewise cosine basis functions to systematically construct symmetry invariant features into the neural network model. We illustrate this DFT2FF approach for amorphous silicon where only 800 DFT configurations are sufficient to achieve an accuracy of 1meV/atom for energy and 0.1eV/Å for forces. We then use the resulting FF model to calculate the thermal conductivity of amorphous Si based on long molecular dynamics simulations. The dramatic speedup in training in our DFT2FF protocol allows the adoption of a new simulation paradigm where an accurate and problem specific FF for a given physics phenomenon is trained on the spot through a quick DFT pre-calculation and FF training.

1. Introduction

Machine learning methods have rapidly developed to solve scientific problems in biology, chemistry, physics and engineering [1-6] in recent years. In the field of atomic and molecular studies, one of the major applications of machine learning is to obtain the quantitative structure-activity relationships (QSAR) [7-9]. In molecular simulations, the relationship between the total energy of a system and its atomic or molecular structure is one of the most important properties because derivatives of the total energy with respect to atomic positions give rise to forces, which can be used to perform molecular dynamic simulations [10, 11].

Such a relationship is described by the potential energy surface (PES) of the system. The PES is difficult to obtain in experiment; rather, it is typically sampled by solving the Schrodinger equation. In practice, density functional theory (DFT) approximation to the Schrodinger Hamiltonian is used. Once the information about the system is calculated using DFT, machine learning models can be applied to fit the PES.

Many PES machine learning models have been developed over the last decade, including generalized symmetry functions [12], smooth overlapping of atomic potentials (SOAP) [13], bag of bonds (BoB) [14], Gaussian approximation potentials (GAP) [15], deep tensor neural networks (DTNN) [16]. Among them, the most widely used model is the high-dimensional neural network potential (HDNNP) model proposed by Behler et al [17]. They employed symmetry functions to collectively map the chemical environment of individual atoms to a set of descriptors (called features), which are subsequently fed into a multilayer neural network to yield the total energy of the system. However, the mapping of the atomic environment to descriptors is not unique, and many choices of symmetry functions have been reported in the literature [12, 18-20].

Compared with the conventional classical molecular force fields, the neural network force field (NNFF) can be more accurate in the atomic configuration space where it is fully trained, but it can fail catastrophically in regions where it is not exposed [21]. One way to make a proper use of this feature of NNFF is to train the NNFF, if not on the flight, but at least on the spot. For each target physical phenomenon, one can first carry out an *ab initio* DFT simulation on a smaller system for a shorter time, while at the same time ensuring that the DFT simulation covers all the possible local atomic configurations essential to the physical phenomena to be studied. This will be followed by a standard and quick NNFF training, and the resulting NNFF can then be used to simulate a much larger system for a much longer time. To make this procedure practical, one has to satisfy the following requirements: (1) a quick generation of large amount of DFT data; (2) a universal NNFF model; (3) a corresponding quick training procedure of this model; and (4) finally, the ability to yield accurate NNFF results compared to DFT data within the desired configurational region.

In this work, we present a new model to fit NNFF which satisfies the above requirements by using atomic energies decomposed from a DFT calculation and piecewise cosines for systematic symmetry features. The model is implemented in TensorFlow® to utilize directed acyclic graph (DAG) to accelerate computation. The resulting NNFF for a test amorphous silicon system has an accuracy of 1.0 meV/atom in energy and 0.1eV/Å for the forces. We also show how the resulting NNFF can be used to calculate material properties which could be too expensive to calculate directly using DFT.

Our work follows closely to the HDNNP model [17]. However, in the current HDNNP approach, typically, ~10,000 DFT trajectory steps will be used to fit the HDNNP models [17, 22]. These can take many days for the DFT calculation. A major advance in the current study is to decompose the DFT total energy of a given system into atomic energies belonging to each atom. Importantly, such atomic energy only depends on the positions of the nearby atoms. As a result, a unified single atom neural network potential (SANNP) model can be used taking into account the data from all atoms. In comparison, in HDNNP, only the total energy of the system is used in the training set although the atomic energies are implicitly assumed in the model. Due to the increase of the data set, we found that ~1000 molecular dynamics (MD) steps are sufficient to train an accurate SANNP. This makes it practical to carry DFT calculation and SANNP training overnight, allowing for an on-the-spot SANNP development.

Combining the energy decomposition method with the piecewise cosine functions model, we show that the training of the SANNP using TensorFlow running on a GPU workstation only takes a few hours (requirement 3) and the resulting SANNP has an accuracy of ~ 1meV/atom for the energy (of a 256 atom system) and 0.1eV/Å for the forces (requirement 4). As an example, we have performed MD simulations to calculate the thermal conductivity of amorphous Si (a-Si), which is difficult to obtain using direct DFT calculations [23]. Although we have calculated Si in different temperatures, from low to melt liquid temperature, and found the procedure equally applicable, in the current study, we will focus on the results of the amorphous Si (a-Si) structure. The a-Si structures can be generated from the random bond

switching model [24], followed by DFT MD simulations at different temperatures. It is also worth noting that the current SANNP can be easily extended to systems with M_e type of atoms. In such cases, each type of center atom will have its own SANNP model, and for a given SANNP model, the number of atom pair features increases by a factor of M_e , while the number of three atom features increases by a factor of M_e^2 . The extension of SANNP to such systems (an ongoing project) further exemplifies the importance of more DFT data, which can be provided by our energy decomposition scheme.

2. Method

2.1 High Dimensional Neural Network Potential (HDNNP)

The HDNNP is a machine learning model developed by Behler et al. [17]. In this model, the total energy of a given system is assumed to be the simple sum of the atomic energies $E_{tot} = \sum_i E_i$. However, the DFT values of E_i are not known, and only E_{tot} is obtained for a given system. Nevertheless, such an assumption allows the construction of a HDNNP model as shown in Figure 1(a), where each horizontal bar is a small multilayer (typically 2 hidden layers) neural network, and the edge weights on the last step connecting E_i to E_{tot} are fixed at 1. More importantly, the neural network parameters for each small network (horizon bar) are the same, such that the whole network consists of N identical smaller networks (for single specie systems), where N is the number of atoms in the system.

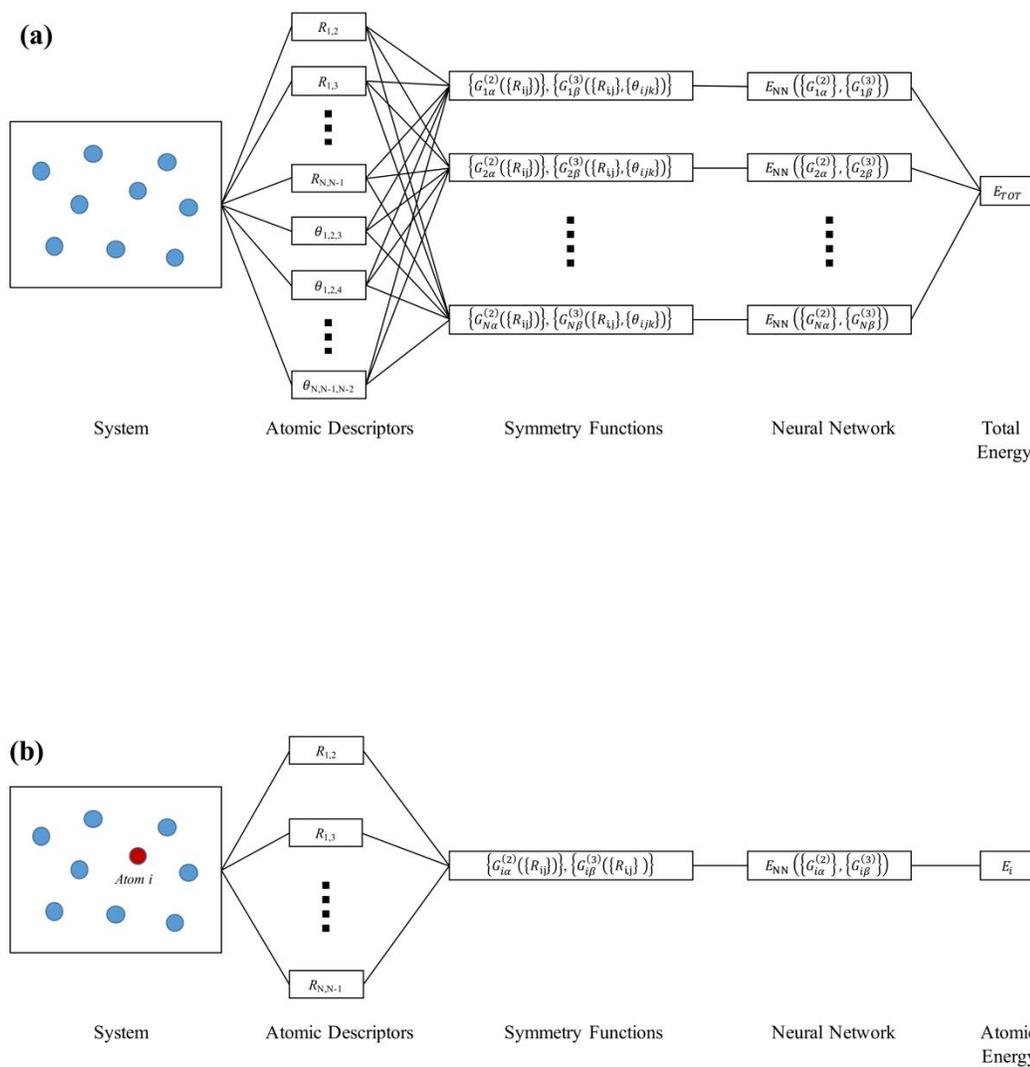


Figure 1. Schematics of the neural network models. (a) High dimensional neural network potential (HDNNP), (b) single atom neural network potential (SANNP). The leftmost box is the simulation system, the second column represents the atomic descriptors including interatomic distances, and three-body angles, the third column represents the symmetry functions, the fourth column represents the neural network for the model, and the fifth column represents the energy term(s) to be trained for.

Mathematically, let $\{\mathbf{p}_{ijj}, \mathbf{q}_{ijk}\}$ be a set of structural descriptors including the interatomic distances (R_{ij}) and angles (θ_{ijk}) between atom j and atom k about atom i , $\phi_\alpha^{(2)}$ and $\phi_\beta^{(3)}$ be the α^{th} 2-body symmetry function and β^{th} 3-body symmetry function, and $E_{NN}(\{G_{i\alpha}\}; \mathbf{w}, \mathbf{b})$ be the multilayer neural network model with descriptors $\{G_{i\alpha}^{(2)}, G_{i\beta}^{(3)}\}$ as its input, and weight parameters \mathbf{w} and bias parameters \mathbf{b} being its fitted model parameters. Then the general mathematical form of the original HDNNP is:

$$G_{i\alpha}^{(2)} = \sum_j \phi_\alpha^{(2)}(\mathbf{p}_{ijj}); \quad G_{i\beta}^{(3)} = \sum_{jk} \phi_\beta^{(3)}(\mathbf{p}_{ijj}, \mathbf{q}_{ijk}) \quad (1)$$

$$E_i = E_{NN}(\{G_{i\alpha}^{(2)}, G_{i\beta}^{(3)}\}; \mathbf{w}, \mathbf{b}) \quad (2)$$

$$E_{\text{tot}} = \sum_i E_i \quad (3)$$

As shown in Equation (1), the summation over atomic index j and j, k enforces permutation invariance, while the proper constructions of $\phi_\alpha^{(2)}$ and $\phi_\beta^{(3)}$ ensure the translational and rotational invariance of the descriptors $G_{i\alpha}^{(2)}$ and $G_{i\beta}^{(3)}$.

2.2 Atomic Energies

As an improvement to energy fitting using the total energy, we propose a way to actually calculate E_i from the DFT calculations. As a result, our network is simplified and consists of only one small network (one bar in Figure 1), thus the dataset is increased by N fold, where N is the number of atoms in the DFT system, which is typically around 100-200.

To expand the dataset for training, the DFT total energy is partitioned into atomic energies outlined in Kang and Wang [23]. The critical point is to rewrite the DFT energy terms (kinetic, electrostatic) as the spatial integration of their respective energy densities [25], such that a Hirshfeld style spatial decomposition can be used to decompose the energy into atomic contributions. More specifically we have:

$$U^{DFT} = \sum_i \left[U_i^{DFT} + E_i^{(NL)} - E_{\text{self}} \right] + \sum_{i \neq j} V^P(|\mathbf{R}_{ij}|)$$

With

$$U_i^{DFT} = \int d\mathbf{r} [t_0(\mathbf{r}) + e_{\text{xc}}(\mathbf{r}) + e_{\text{CC}}(\mathbf{r})] \frac{w_i(|\mathbf{r}-\mathbf{R}_i|)}{\sum_j w_i(|\mathbf{r}-\mathbf{R}_j|)} \quad (4)$$

Where $w_i(r)$ is the radial charge density function [26] of the neutral atom at site i , $t_0(\mathbf{r})$ is the electronic kinetic energy density, $e_{\text{xc}}(\mathbf{r})$ is the exchange-correction energy density, $e_{\text{CC}}(\mathbf{r})$ is the Hartree energy density, $E_i^{(NL)}$ is the nonlocal contribution from the pseudopotential for atom i , E_{self} is an onsite energy integral constant, and $V^P(|\mathbf{R}_{ij}|)$ is the Coulomb interaction between ion pairs. Summarizing, individual atomic energy E_i has the following expression:

$$E_i^{DFT} = U_i^{DFT} + E_i^{(NL)} - E_{\text{self}} + 1/2 \sum_{j \neq i} V^P(|\mathbf{R}_{ij}|) \quad (5)$$

It is noted that the above energy decomposition is not unique. But for our SANNP development, this is not critical, as long as the sum of E_i agrees with the whole system total energy, and E_i is a local property that only depends on the atomic configuration near atom i . It is known that there are remaining challenges for NNFF when long range Coulomb interaction is strong. In such cases, atomic charges might need to be fitted [27], and the corresponding energy contribution needs to be subtracted before applying the above decomposition scheme. As shown in Supplementary Materials S1, for non-polarized systems like amorphous Si, the atomic energy E_i is indeed a local property of its atomic configuration.

2.3 Piecewise cosine symmetry functions

In order to use an artificial neural network to fit the atomic or total energies, the surrounding chemical environment of each atom has to be mapped to a set of descriptors using symmetry invariant functions. To capture the complicated correlation within such an environment, symmetry functions involving two body pairs and three atom triplets are used, as described in Equation (1). In addition to the Gaussian-like symmetry functions, other approaches have been developed in literature, including Zernike [18], Bispectrum [19], and Chebyshev radial

distributions [20]. However, the resulting atomic descriptors are typically nontrivial functions of the atomic coordinates, which makes the derivative with respect to the atomic position complicated. In general, there are two approaches to construct the descriptors or features. One is to use some spectrum functions (e.g., the Bispectrum and Chebyshev), which are smooth and might have some transferability in real space, and the other one is to use spatially localized functions that are invariant with respect to translation, rotation and permutation. We have adopted the later approach.

One example of local representations in numerical calculation is the piecewise linear functions in finite element analysis. These piecewise linear functions are defined on a set of nodes, such that any continuous function defined on this domain can be approximated as linear combinations of these local functions. For our purpose, in order to calculate the force, the derivative of the piecewise function with respect to the atomic position is needed. As a result, the piecewise linear shape functions are modified to differentiable piecewise cosine functions for our neural network model. The shapes of these piecewise cosine functions are shown in Figure 2 (a).

Using the piecewise cosine functions, the symmetry functions can be constructed in simple forms as follows:

$$R_{\alpha}^k = R_{inner} + (\alpha - 1)h_k \quad \text{where } \alpha = 1, 2, \dots, M_k$$

$$\varphi_{\alpha}^{(k)}(R_{ml}) = \begin{cases} \frac{1}{2} \cos\left(\frac{R_{ml}-R_{\alpha}^k}{h_k} \pi\right) + \frac{1}{2}, & |R_{ml} - R_{\alpha}^k| < h_k \\ 0, & \text{Otherwise} \end{cases} \quad (6)$$

$$G_{\alpha,l}^{(2)} = \sum_m \varphi_{\alpha}^{(2)}(R_{ml}) \quad (7)$$

$$G_{\alpha\beta\gamma,l}^{(3)} = \sum_{m,n} \varphi_{\alpha}^{(3)}(R_{ml}) \varphi_{\beta}^{(3)}(R_{nl}) \varphi_{\gamma}^{(3)}(R_{mn}) \quad (8)$$

where:

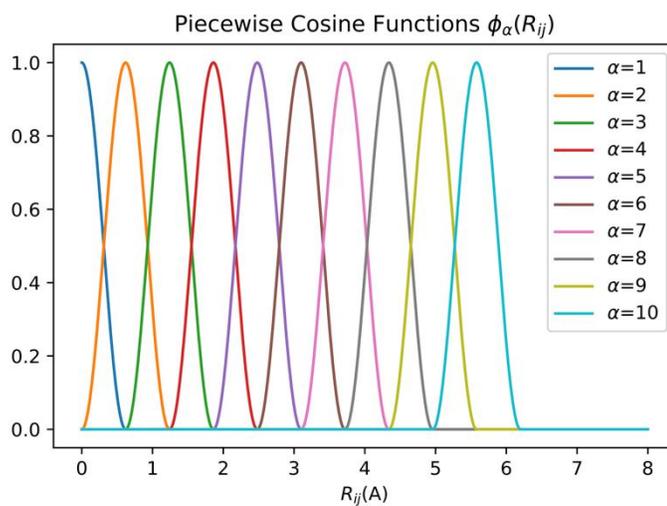
- $k=2,3$ specifies the 2-body and 3-body terms respectively,

- $h_k = (R_{\text{outer}} - R_{\text{inner}})/M_k$ is the width of the piecewise cosine function, with M_2 being the number of piecewise cosine basis functions for the 2-body term, and M_3 being the number of basis functions for each side of the 3-body term,
- $R_{ml} = |R_l - R_m|$ are the interatomic distance between atom l and atom m , and
- Σ_m and $\Sigma_{m,n}$ are sums over all atoms within the R_{outer} cutoff of atom l .

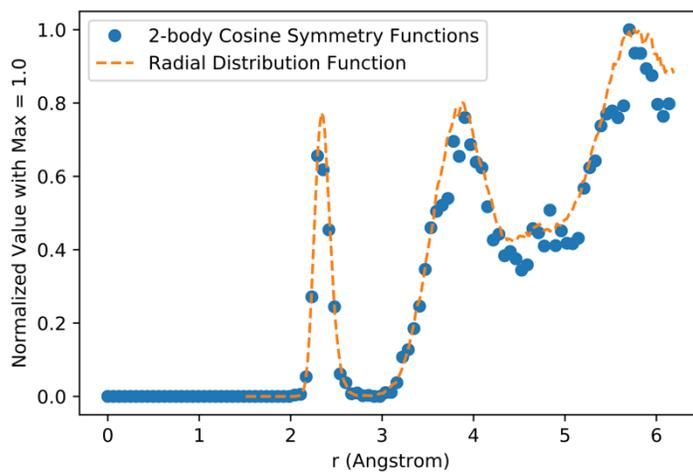
Once the inner and outer cutoff radii are determined, the two-body symmetry functions are completely determined by a single number M_2 . However, it is not practical to set $M_3=M_2$ for the number of three-body cosine basis functions because the number of three body symmetry functions will scale as $O(M_3^3)$. Therefore, we have used a balanced set of symmetry functions characterized by a single number M , where $M_2 = M^2$ and $M_3 = M$.

Our symmetry functions have similar forms as the Gaussian symmetry functions used in HDNNP [17] for the 2-body term, but with important simplifications. First, we no longer need to multiply the symmetry function by an arbitrary decay function to ensure that the symmetry function goes to 0 smoothly at R_{outer} (and also at R_{inner} if it is not 0 already), because these symmetry functions are local and the rightmost function already decays to 0 at R_{outer} . In a way, these functions are more local than the Gaussian functions since they go to zero outside their perspective ranges. Moreover, in the three-body term, all three sides of each atomic triplet are treated equally, and an arbitrary cosine term is no longer needed to describe the angle dependence. These piecewise cosine functions are shown in Figure 2(a). Physically, the values of the $G_{\alpha,l}^{(2)}$ simply represent the pair correlation function for the atom l with α being the distance from the center atom l . This is shown in Figure 2(b). Note, if we have an infinitely large number of M_2 and M_3 (infinitely localized functions), one can show that, if all the $G_{\alpha,l}^{(2)}$ and $G_{\alpha\beta\gamma,l}^{(3)}$ are determined, the local atomic positions within R_{outer} will be completely determined (up to the translation, rotation and permutation degree of freedoms). Thus $G_{\alpha,l}^{(2)}$ and $G_{\alpha\beta\gamma,l}^{(3)}$ in combination is a complete descriptor of the local atomic

configuration for atom l . Should there be different atom types, the α, β, γ index should also include the information of atom types, besides the distance $h_{(\alpha,\beta,\gamma)}$.



(a)



(b)

Figure 2. Piecwise cosine functions as basis functions to construct the symmetry functions. (a) graphical representation of the piecwise cosine functions when $M=12$ basis functions

are used. (b) comparison of the radial distribution and the normalized values of the two-body piecewise cosine features with their respective nodes with $M=100$.

In summary, the original HDNNP formulation is modified in two major ways. First, since the total energy is decomposed to atomic energies, the neural network will be trained against the individual atomic energies directly. Second, the piecewise cosine functions are used as basis functions constructing symmetry functions to obtain the descriptors of the chemical environment.

2.4 Training procedure

In traditional classical force fields, although the functional form of the energy contributions or forces is known, the number of corresponding parameters is small and difficult to train because different force fields are based on different functional forms. Thus, classical force fields are usually trained with chemical and physical intuition on carefully selected quantum mechanical trajectories. On the other hand, neural network based machine learning models have many more parameters to flexibly fit any configurations, and the fitting is made possible by the back-propagation procedure [28]. Thus, an efficient training procedure is crucial to realize the on-the-spot SANNP development strategy.

Since our dataset contains both atomic energies and forces, a multi-step procedure is used to train the neural network potential more effectively. Because of the large amount of atomic energies obtained from energy decomposition, and the straightforward back-propagation algorithm, the neural network is first trained with the atomic energies using stochastic gradient descents. Since the atomic descriptors are used thousands of times as the training loops through the dataset repeatedly, it is much more preferable to store the descriptors obtained from the piecewise cosine symmetry functions on disk to avoid wasteful repeating calculations. Once the model is finished pre-training with energies, forces can be added to further adjust the parameters of the neural network for a smooth fitting of the PES.

For our amorphous silicon system with 256 atoms, a large set of features can be stored directly in memory to avoid excessive I/O on the hard drive. Therefore, initially, atomic

descriptors of a 100 randomly chosen DFT configurations along the MD trajectory, which correspond to 25,600 atomic energies, are calculated and stored in memory to pre-converge the parameters of the artificial neural network. However, to avoid overfitting, the artificial neural network is only trained 5000 iterations with these 100 configurations. After the initial training, all the atomic energies in the training set are used to train the neural network for another 500 iterations.

After the neural network potential has been trained against the DFT energies E_i , DFT forces can be used to fine-tune the neural network to obtain accurate forces. For each trajectory, we can define a loss function that includes both atomic energies and atomic forces as $L = \text{MSE}(\{E_i\}) + \alpha \times \text{MSE}(\{\mathbf{F}_i\})$, where the SANNP forces on each atom are obtained by analytical differentiation of the total neural network energy with respect to the atomic coordinates, and $\text{MSE}()$ is the mean squared error. The training using L with respect to the neural network parameters \mathbf{w} and \mathbf{b} can be carried out using the TensorFlow® library. For this study, a two-hidden-layer neural network model with 500 nodes in each layer is used. The Adam optimizer [29] with a learning rate of 0.0001 is used to minimize the MSE loss function.

At last, to perform MD simulations at a certain temperature T_o , the neural network training must be supplemented with a higher temperature simulation data to cover a larger area of the configurational space. Thus, after training the neural network with DFT trajectories from the target temperature, higher temperature DFT trajectories are included to train for another 100 iterations for the combined training set.

The above model is implemented in our publicly accessible custom code [Supplementary Materials S2]. Similar to quantum mechanical calculations in which near-complete basis sets are used when comparing different methods, a large two-layer neural network with 500 nodes each is used here to avoid the finite size effects of the neural network on the SANNP method. As shown in [Supplementary Materials S3], the neural network potential is well converged using 500 nodes, and the error is comparable to the case with only 40 nodes. The training set has 800 DFT configurations, each with 256 silicon atoms. The piecewise cosine functions

with $M=10$, or equivalently 550 features in which $M_2 = M^2$ and $M_3 = M$, are used. The calculations are performed on an NVIDIA Titan X GPU. Each training iteration through the whole training set takes about 4.5 minutes when both atomic energies and forces are used. The forces converge to 0.1 eV/Å in the validation set after 150 training iterations, which takes about 11.4 GPU hours. When $M=5$, or equivalently 100 features, is used, it takes 2.4 minutes per training iteration. The forces converge to 0.13 eV/Å in 200 iterations, which correspond to about 8.1 hours.

To compare the accuracy and training speed between the SANNP and the HDNNP approaches, we construct an HDNNP model with the same set of features as our SANNP model and train them with the same data, but with HDNNP training on the total energies, E_{tot} , and with SANNP training on atomic energies, E_i . Mathematically, if the training is long enough, the HDNNP will give the same result as SANNP. However, because the models are trained on different data, the convergence will be different. Using the same procedure and after training for 200 iterations using forces, the HDNNP results in an error of 0.17 eV/Å, and the SANNP results in an error of 0.13 eV/Å. The reason of the larger error of HDNNP is due to a much slower convergence, since the HDNNP over-fits on the 800 total energies, while the SANNP is trained appropriately with $800 \times 256 = 204600$ atomic energies.

3. Results and Discussion

3.1 Comparison between DFT and neural network potential

To validate our SANNP model using energy decomposition, piecewise cosine symmetry functions, and the proposed training procedure, we train the corresponding neural network for a periodic system of amorphous silicon, which is initially generated with a random covalent band switching model [24]. A set of 1000 DFT configurations is obtained from an *ab initio* molecular dynamics simulation of a periodic box with 256 silicon atoms [Supplementary Materials S4], such that 800 points are used for training, 100 points are used for validation, and 100 points are used for testing. By decomposing the total energy into individual energies, the training set contains 204,800 atomic energies and 614,400 atomic

forces. After training with the procedure described above, the test set errors in atomic energies, total energies, and forces are shown in Figure 3.

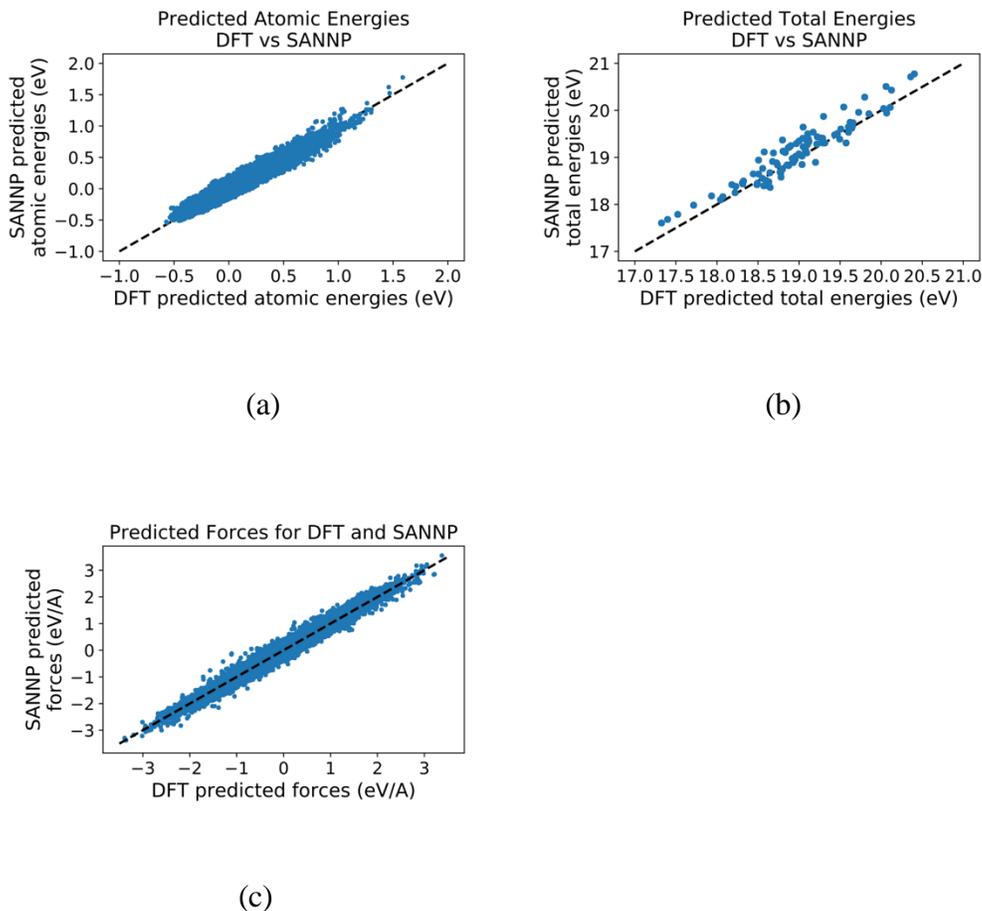


Figure 3. Comparison between the fitted SANNP model and DFT for the test set with 100 configurations. (a) comparison of SANNP atomic energies and DFT atomic energies in the test set. (b) comparison of the SANNP total energies and DFT total energies in the test set. (c) comparison of the SANNP forces and DFT forces in the test set. Note that none of the test set data are used in training the SANNP model.

Overall, the neural network potential is in good agreement with the DFT results, with the RMSE's in $\{E_i\}$, E_{tot} and $\{\mathbf{F}_i\}$ being around 50meV, 1.0meV/atom, and 0.10eV/A respectively in the test set. It is important to note that the RMSE per atom for E_{tot} is $N^{-1/2}$ of the RMSE for E_i due to error cancellations. It should also be noted that the neural network was not trained directly against E_{tot} , and yet an extremely small E_{tot} RMSE of 1.0meV/atom

was recovered. All these indicate that our new approach using atomic energies E_i can achieve high accuracy with a small number of DFT trajectories.

3.2 Application of the cosine-based symmetry functions

To evaluate the effectiveness of the piecewise cosine localization model, we have compared our model with the Gaussian-like symmetry functions used in the original HDNNP model. Except the change from the cosine like function to Gaussian function, all the other procedures are the same (however, only E_i is used in the training, atomic force F_i is not used in this test). As shown in Table 1, when training against the atomic energies, the piecewise functions with $M=4$ (or a total of 56 feature functions $M_2 = M^2$ and $M_3 = M$) achieve a similar accuracy as the case with more than 100 Gaussian-like symmetry functions. When more piecewise cosine functions with $M=5$ or a total of 100 functions are used, the energies can be fitted even better.

Basis sets (Total number of functions)	RMSE of E_i on the test set	MAE of E_i on the test set
Gaussian (110)	54.8 meV	43.5 meV
$M=4$ cosine (56)	57.7 meV	45.7 meV
$M=5$ cosine (100)	45.9 meV	36.1 meV

Table 1. Comparison of the symmetry functions for the training atomic energy E_i .

In fact, the quality of the piecewise cosine functions can be systematically improved by adjusting one parameter, M , with the number of two-body symmetry functions is $M_2 = M^2$, and the number of three-body symmetry functions is $(M^2 + M^3)/2$. As shown in Figure 4, both the energies and forces are converging towards certain limits as M increases. In addition, the quality of the basis set can also be adjusted by changing the inner cutoff distance. In the case of amorphous silicon, the interatomic distance between any two Si atoms is rarely less than 1.9Å. By increasing this inner cutoff, the piecewise cosine functions are more concentrated in the region of interest, thus allowing a better description and resulting in a more accurate model. As shown in Figure 4, the neural network potential with an inner cutoff of 1.9Å is consistently better than inner cutoff of 0Å. In summary, the piecewise cosine functions can

be adjusted by tuning the number of functions (M) and the range of interatomic distances, which are both intuitive parameters to be adjusted in practice.

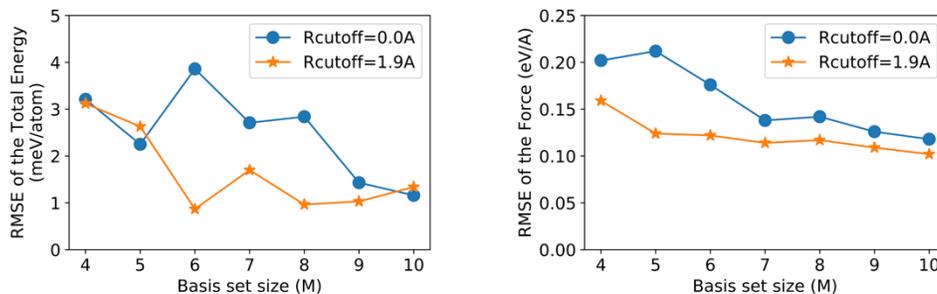


Figure 4. Errors in (a) total energies and (b) forces as functions of the size M of the piecewise cosine functions. Inner cutoffs of $R_{\text{inner}}=0.0\text{\AA}$ (or no cutoff) and $R_{\text{inner}}=1.9\text{\AA}$ are also compared.

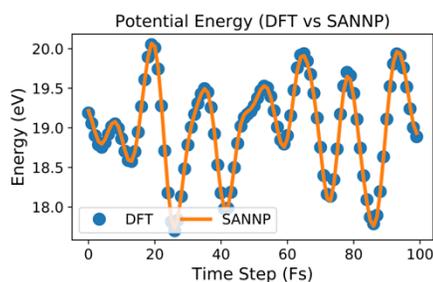
3.3 Comparison of DFT and NN for MD

One of our main goals is to use the SANNP to perform molecular dynamic simulation. In addition to having small errors in energies and forces in the test set, an accurate neural network potential must also be able to reproduce similar quantities along its MD trajectory. In particular, the smoothness of the potential as well as the atomic forces is important. For this test, we have performed an NVE molecular dynamics simulation on another amorphous silicon structure (not in the training set) using DFT for 1ps, and a random trajectory interval of 100 fs is chosen to compare the energies and forces between DFT and the previously trained SANNP. As shown in Figure 5(a), along the AIMD trajectories, the energies between DFT and SANNP match almost perfectly, with an RMSE of total energy of 1.1meV/atom, which is only slightly higher than that of the test set. Since the forces at each trajectory include all components of all atoms, to compare the DFT forces and the SANNP forces, we have projected the forces along the MD trajectory directions, as shown in Figure 5(b), and have also calculated the unit vector dot product between the DFT and SANNP forces, as shown in Figure 5(c). Overall, the projected forces agree well between DFT and SANNP with an RMSE of 0.13eV/A, and the scaled dot product indicates that the SANNP forces recover almost 99% of the DFT forces throughout the trajectory. Such a near unity dot-

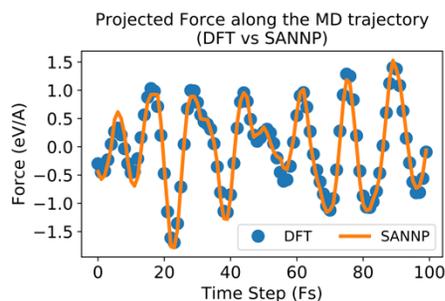
product means the forces for almost all atoms are in the same directions between SANNP and DFT.

Since silicon is an important material, many empirical force fields have been developed. The energies for these classical force fields can also be calculated along the above AIMD trajectories. As shown in Figure 5(d), energies obtained for all methods have the same overall trend and local extrema, but the variation in energy at each trajectory is quite large for different force fields, as well as to the DFT energies. This indicates a much superior SANNP accuracy compared to other classical force fields, if the DFT energy is used as the reference.

The SANNP can also be compared to DFT with different practical parameters (pseudopotentials and exchange-correlation functionals). As shown in Figure 5(e), the energies for different DFT are calculated along the same AIMD trajectories, and they are relatively close to each other. By zooming in onto the first 10 time steps of the trajectories, the differences between different DFT runs can be shown more clearly. As shown in Figure 5(f), the SANNP follows the original DFT results very closely, while different levels of DFT give much bigger errors. This indicates that the SANNP is already within the errors between different choices of DFT.



(a)



(b)

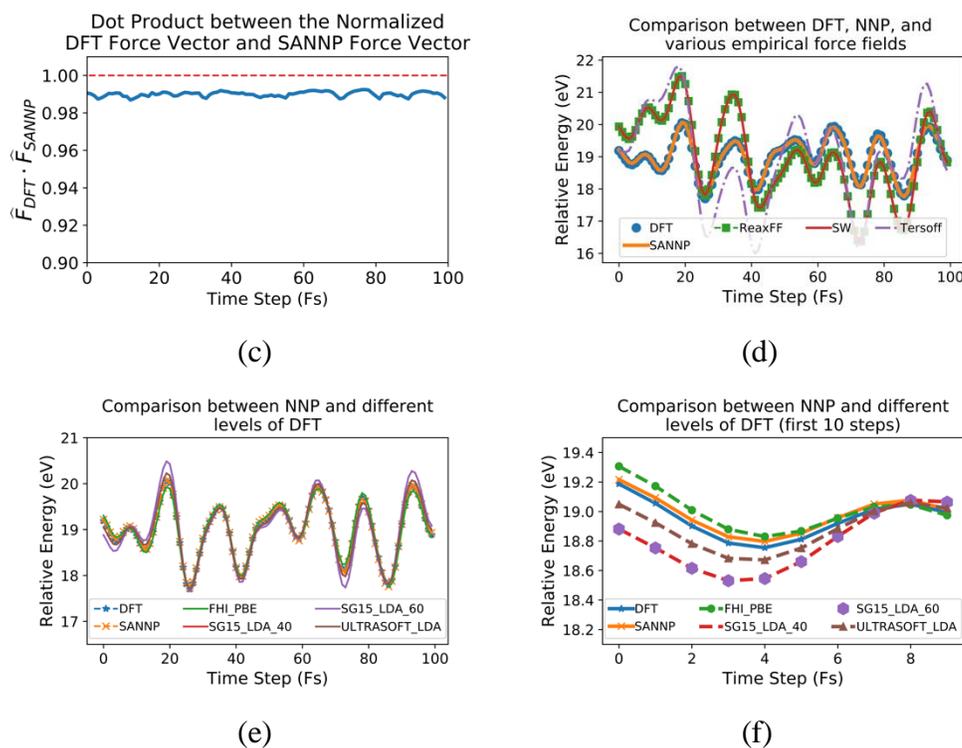


Figure 5. Comparison of DFT, neural network potential, and empirical force fields along a sequence of AIMD trajectories. (a) total energies between DFT and NNP (b) projected forces between DFT and neural network forces (c) scaled dot product between DFT and neural network forces (d) total energies between DFT, NNP, and classical force fields (with constant shifts for force field energies); (e) total energies between the neural network potential and various levels of DFT over 100fs (e) total energies between the neural network potential and various levels of DFT over 10fs

Although the neural network fits the DFT trajectory well, it is still necessary to compare the dynamics between DFT and SANNP to ensure that similar structural properties are generated. Starting from the same structure and initial velocities, NVE simulations are performed using DFT and SANNP independently. After 1ps, the radial distribution function and the normalized angular distribution are compared. The radial distribution is obtained from the total bond distribution normalized by r^2 , the angular distribution is normalized by $\sin(\theta)$. As shown in Figure 6, both distributions are almost indistinguishable between DFT and SANNP, confirming that the SANNP is capable of reproducing the DFT results through MD simulations.

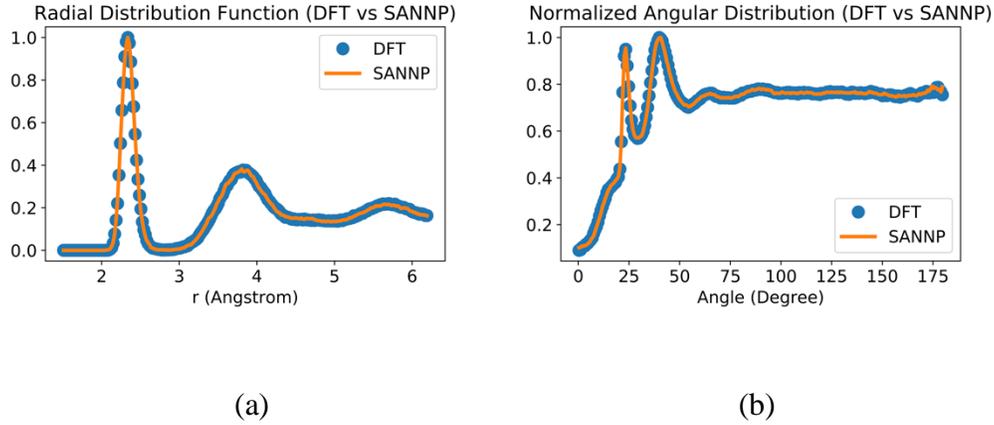


Figure 6. Comparison of (a) radial distributions and (b) angular distributions between DFT (blue) and single atom neural network potential, SANNP (orange).

3.4 Application of the NNP to thermal conductivity calculations

As shown in the previous sections, the SANNP is capable of performing MD simulations with DFT level of accuracy. Since the SANNP is significantly less computationally intensive than DFT, long time and large-scale simulations can be performed. One class of problems that is difficult to be calculated with DFT level of accuracy is the classical transport properties, e.g., the thermal conductivity of amorphous silicon. By employing the SANNP to perform molecular dynamics simulations over a long time scale, the heat current auto-correlation function (HCACF), $\langle \mathbf{J}(t)\mathbf{J}(0) \rangle$, can be obtained. Using the Green-Kudo formulation, the thermal conductivity equals the integration of the HCACF,

$$\kappa = \int^t \langle \mathbf{J}(t)\mathbf{J}(0) \rangle dt \quad (9)$$

Although there is no unique spatial origin of \mathbf{R}_i to calculate the heat current $\mathbf{J}(t) = \frac{d}{dt} \sum_i \mathbf{R}_i E_i$, an alternative formulation [30] can be used to avoid this non-uniqueness as long as the atomic energies are explicitly represented as a function of the atomic coordinates, as in the SANNP. For this, we have:

$$J(t) = \sum_i \mathbf{v}_i \left(\frac{1}{2} m_i v_i^2 + E_i \right) + \sum_i \sum_{j \neq i} \mathbf{R}_{ij} \nabla_i E_j \cdot \mathbf{v}_i \quad (10)$$

Where E_i is the atomic energy from the SANNP.

To obtain the HCACF by taking ensemble average of the heat current in Equation (10), we first shift the temperature of the system to the target temperature using the Andersen thermostat such that the canonical ensemble is correctly sampled. After the system is further stabilized at the target temperature for another 1ps, an NVE simulation is performed to sample the dynamics and obtain $J(t)$.

The HCACF for amorphous silicon at 300K is shown in Figure 7(a), in which most of the correlation is within the first few hundred femtoseconds, and it quickly decays to zero. By integrating over the HCACF, the corresponding thermal conductivity κ is integrated to be 1.59 ± 0.1 W/m·K, as shown in Figure 7(b). This can be compared with the direct DFT calculated value of 1.4 ± 0.3 W/m·k [23] at the same temperature, as well as the experimental range of 1 to 2 W/m·K [31-33]. We have also calculated the temperature dependence of κ as shown in Figure 7(c). We see that κ decreases in a power law rate as temperature increases, the same trend as found in crystal Si structure [34]. However, the rate of decreasing is much slower in our case, a result of the randomness in the amorphous Si, where the phonon scattering is caused mostly by the structure randomness instead of by the temperature dependent nonharmonic scattering. As a result, the temperature dependence is much smaller.

The accurate computation of the thermal conductivity using the SANNP implies that the heat current $J(t) = \frac{d}{dt} \sum_i \mathbf{R}_i E_i$, and thus the local atomic energies, are properly obtained using our atomic energy decomposition scheme and accurately trained using our neural network model.

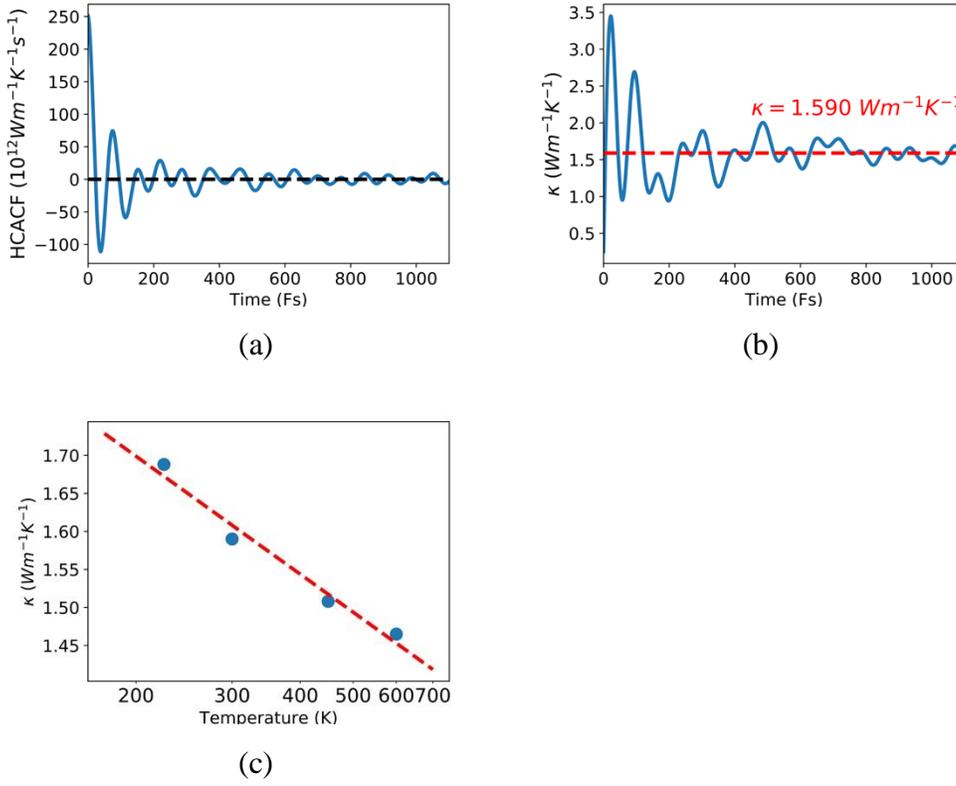


Figure 7. (a) The time evolution of the heat current auto-correction function (HCACF), $\langle J_x(t)J_x(0) \rangle$, for amorphous silicon at 300K; (b) The time integration of the HCACF using Equation (9) at 300K; (c) The temperature dependence of the thermal conductivity from 150K to 600K.

4. Conclusion

In this study, we developed a machine learned based scheme to partition the DFT total energy into atomic energies that depend only on the atoms nearby to a given atom. This leads to a unified single atom neural network potential (SANNP) model that uses the data for all atoms in training this SANNP. Compared to the traditional HDNNP method using only the total energy of the system, SANNP acquires hundreds of times more energy information from the same DFT. As a result, we found that only 1000 MD steps (which takes about half day to finish using GPU by the PWmat code [35-37]) are sufficient to train an accurate SANNP, which dramatically reduces the training time while dramatically increasing the accuracy.

In addition, we have deployed a new universal set of symmetry invariant feature functions using local piecewise cosine basis. We show that using piecewise cosine functions to construct the symmetry features provides a systematic and mathematically efficient way to represent the atomic configuration of nearby atoms for a given central atom. This provides a universal model applicable to any system. Combining the energy decomposition method with the piecewise cosine functions model, we show that the training of the SANNP using TensorFlow running on GPU workstation takes 1/10 the time of standard methods while attaining total energy accuracy of $\sim 1\text{meV/atom}$ (of a 256 atom system) and force accuracy of 0.1eV/\AA . This accuracy is comparable, or even higher, than the uncertainties from using different pseudopotentials in DFT calculations. We anticipate that the dramatic reduction in the number of DFT pre-calculations (with smaller number of steps) with the dramatic decrease in NN training time, plus the high accuracy and systematics of the SANNP model, will enable DFT accuracy for large scale simulations.

References

1. M. I. Jordan and T.M. Mitchell, *Machine learning: Trends, perspectives, and prospects*, Science **349**(6245), 255-260 (2015).
2. M. W. Libbrecht and W. S. Noble, *Machine learning applications in genetics and genomics*, Nat. Rev. Genet. **16**(6): 321-332 (2015).
3. C. Angermueller, T. Pärmamaa, L. Parts, and O. Stegle, *Deep learning for computational biology*, Molecular Systems Biology **12**(7): 878 (2016).
4. R. Ramprasad, R. Batra, G. Piliya, A. Mannodi-Kanakkithodi, and C. Kim, *Machine learning in materials informatics: Recent applications and prospects*, npj Computational Materials **3**, 54 (2017).
5. P. Baldi, P. Sadowski, and D. Whiteson, *Searching for exotic particles in high-energy physics with deep learning*, Nature Communications **5**, 4308 (2014).

6. J. Carrasquilla and R. G. Melko, *Machine learning phases of matter*, Nature Physics **13**, 431-434 (2017)
7. R. D. King, S. H. Muggleton, A. Srinivasan, and M. J. Sternberg, *Structure-activity relationships derived by machine learning: the use of atoms and their bond connectivities to predict mutagenicity by inductive logic programming*, PNAS **93**(1) 438-442 (1996)
8. A. Z. Dudek, T. Arodz, J. Galvez, *Computational methods in developing quantitative structure-activity relationships (QSAR): a review*, Combinatorial Chemistry & High Throughput Screening **9**(3), pp. 213-228(16) (2006).
9. J. Ma, R. P. Sheridan, A. Liaw, G. E. Dahl, and V. Svetnik, *Deep neural nets as a method for quantitative structure-activity relationships*, J. Chem. Inf. Model. **55**(2), pp 263-274 (2015).
10. Z. Li, J. R. Kermode, and A. De Vita, *Molecular Dynamics with On-the-Fly Machine Learning of Quantum-Mechanical Forces*, Phys. Rev. Lett. **114**, 096405 (2015)
11. S. Chmiela, A. Tkatchenko, H. E. Sauceda, I. Poltavsky, K. T. Schütt, and K.-R. Müller, *Machine learning of accurate energy-conserving molecular force fields*, Science Advances **3**(5), e1603015 (2017).
12. J. Behler, *Atom-centered symmetry functions for constructing high-dimensional neural network potentials*, J. Chem. Phys. **134**, 074106 (2011).
13. A. P. Bartók, R. Kondor, and G. Csányi, *On representing chemical environments*, Phys. Rev. B. **87**, 184115 (2013)
14. K. Hansen, F. Biegler, R. Ramakrishnan, W. Pronobis, O. A. von Lilienfeld, K.-R. Müller, and A. Tkatchenko, *Machine learning predictions of molecular properties: Accurate many-body potentials and nonlocality in chemical space*, J. Phys. Chem. Lett. **6** (12), pp 2326-2331 (2015).

15. V. L., Deringer and G. Csányi, *Machine learning based interatomic potential for amorphous carbon*, Phys. Rev. B. **95**, 094203. (2017)
16. K. T. Schütt, F. Arbabzadah, S. Chmiela, K.-R. Müller, A. Tkatchenko, *Quantum-chemical insights from deep tensor neural networks*, Nature Communications **8**, 13890 (2017).
17. J. Behler and M. Parrinello, *Generalized neural-network representation of high-dimensional potential-energy surfaces*, Phys. Rev. Lett. **98**, 146401 (2007).
18. M. Novotni and R. Klein, *Shape retrieval using 3D Zernike descriptors*, Computer Aided Design **36**(11), 1047-1062 (2004).
19. A. P. Bartók, M. C. Payne, R. Kondor, and G. Csányi, *Gaussian Approximation Potentials: The Accuracy of Quantum Mechanics, without the Electrons*, Phys. Rev. Lett. **104**, 136403 (2010).
20. N. Artrith, A. Urban, and G. Ceder, *Efficient and accurate machine-learning interpolation of atomic energies in compositions with many species*, Phys. Rev. B **96**, 014112 (2017).
21. J. Behler, *First Principles Neural Network Potentials for Reactive Simulations of Large Molecular and Condensed Systems*, Angewandte Chemie - International Edition **56**(42), 12828-12840, (2017).
22. J. R. Boes, M. C. Groenenboom, J. A. Keith, J. R. Kitchin, *Neural network and ReaxFF comparison for Au properties*. International Journal of Quantum Chemistry, **116**(13), 979–987. (2016)
23. J. Kang and L.-W. Wang, *First-principles Green-Kubo method for thermal conductivity calculations*, Phys. Rev. B **96**, 020302(R) (2017).

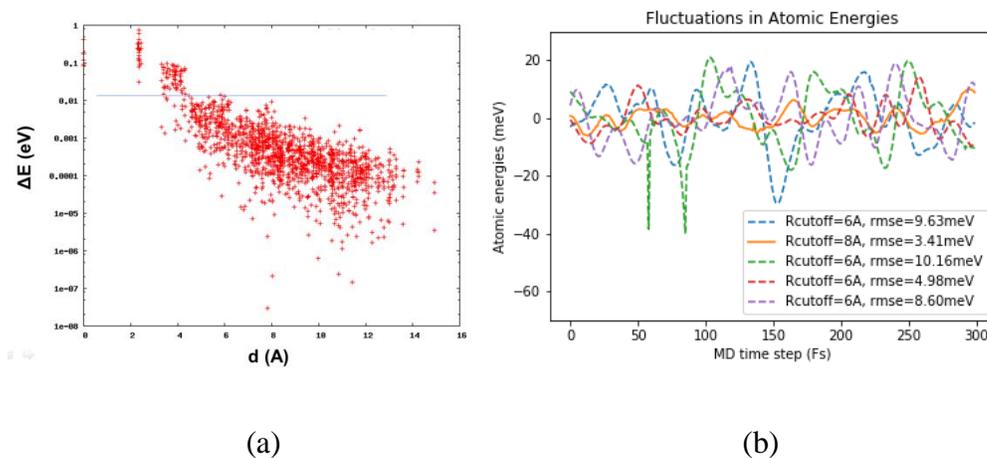
24. F. Zheng, H. H. Pham, and L.-W. Wang, *Effects of the c-Si/a-SiO₂ interfacial atomic structure on its band alignment: an ab initio study*. Phys. Chem. Chem. Phys. **19**, 32617–32625 (2017)
25. M. Yu, D. R. Trinkle, and R. M. Martin, *Energy density in density functional theory: Application to crystalline defects and surfaces*, Phys. Rev. B **83**, 115113 (2011).
26. L.-W. Wang, *Charge-Density Patching Method for Unconventional Semiconductor Binary Systems*, Phys. Rev. Lett. **88**, 256402 (2002).
27. N. Artrith, T. Morawietz, and J. Behler, *High-dimensional neural-network potentials for multicomponent systems: Applications to zinc oxide*, Phys. Rev. B **83**, 153101 (2012).
28. D. E. Rumelhart, G. E. Hinton, and R. J. Williams, *Learning representations by back-propagating errors*, Nature **323**, 533–536 (1986).
29. D. P. Kingma and J. Ba, *Adam: A Method for Stochastic Optimization*, arXiv:1412.6980v9.
30. Z. Fan, L. F. C. Pereira, H.-Q. Wang, J.-C. Zheng, D. Donadio, and A. Harju, *Force and heat current formulas for many-body potentials in molecular dynamics simulations with applications to thermal conductivity calculations*, Phys. Rev. B **92**, 094301(2015).
31. D.G. Cahill, H.E. Fischer, T. Klitsner, E. T. Swartz, and R. O. Pohl, *Thermal conductivity of thin films: Measurements and understanding*, Journal of Vacuum Science & Technology A: Vacuum, Surfaces, and Films, **7**(3), 1259–1266 (1989).
32. H. Wada and T. Kamijoh, *Thermal Conductivity of Amorphous Silicon*, Jpn. J. Appl. Phys. **35** L648 (1996).
33. B. L. Zink, R. Pietri, and F. Hellman, *Thermal conductivity and specific heat of thin-film amorphous silicon*, Phys. Rev. Lett. **96**, 055902 (2006).

34. C. Carbogno, R. Ramprasad, and M. Scheffler, *Ab Initio Green-Kubo Approach for the Thermal Conductivity of Solids*, Phys. Rev. Lett. **118**, 175901 (2017).
35. Long Xun Kuang Teng Inc: <http://www.pwmat.com>
36. W. Jia, J. Fu, Z. Cao, L. Wang, X. Chi, W. Gao, L.-W. Wang, *Fast plane wave density functional theory molecular dynamics calculations on multi-GPU machines*, J. Comput. Phys. **251**, 102-115 (2013).
37. W. Jia, Z. Cao, L. Wang, J. Fu, X. Chi, W. Gao, and L.-W. Wang, *The analysis of a plane wave pseudopotential density functional theory code on a GPU machine*. Computer Physics Communications **184**(1), 9-18. (2013)

Supporting Information

S1. SANNP code and Training data

Figure S1(a) shows how neighboring E_i changes when a center atom j has a displacement. One can see that the change of E_i decays exponentially with the distance $R_{ij} = |\mathbf{R}_j - \mathbf{R}_i|$. Figure S1(b) shows how the atomic energy E_i of a center atom changes when the nearby atoms within a R_{cutoff} are fixed while the outside atoms are moving in a MD simulation at 300K. Several center atoms i are chosen for the tests, and different R_{cutoff} are also examined. When $R_{\text{cutoff}}=6\text{\AA}$, there are roughly 60 atoms within the cutoff radius, and the variation in E_i due to the motions of the outside atoms has a RMSE of 10 meV or less. This will thus be the upper limit of our SANNP fitting when $R_{\text{cutoff}}=6\text{\AA}$ is used. The analyses of E_i here provide another advantage of our SANNP method: the upper limit of accuracy can be obtained prior to the training of SANNP, thus R_{cutoff} can be chosen based on the target accuracy.



(a)

(b)

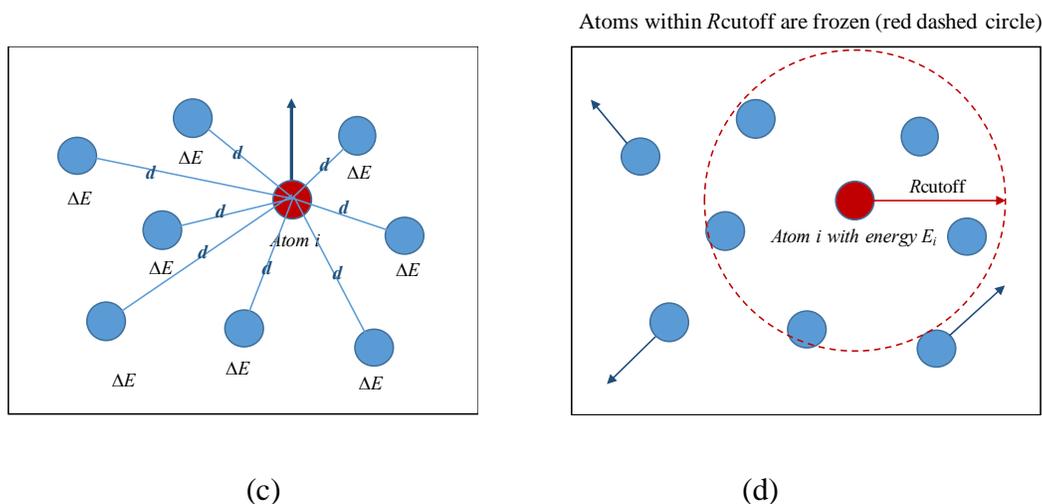


Figure S1. Fluctuation in the employed atomic energy E_i . (a) change of energy of atom j when a chosen atom i displaces. (b) change of energy of atom i with its neighboring atoms within R_{cutoff} frozen and the surrounding atoms outside of R_{cutoff} moving according to an NVE ensemble. (c) schematics of atomic motions for (a). (d) schematics of atomic motions for (b).

S2. SANNP code and Training data

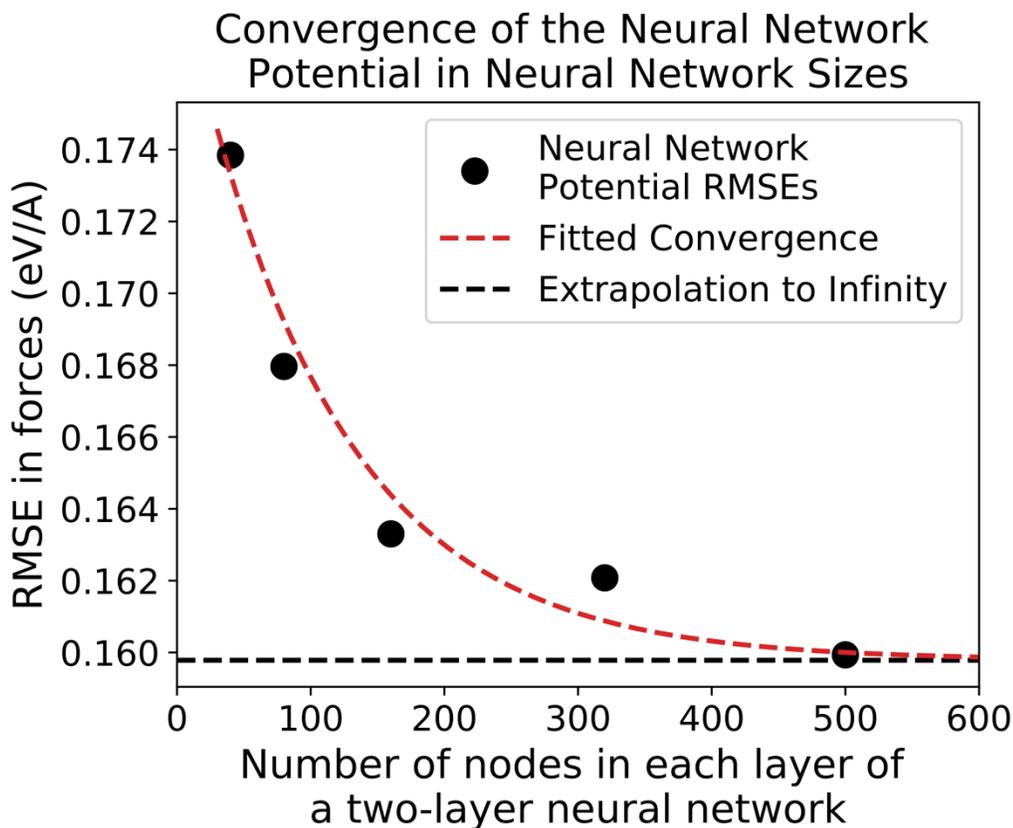
We have written a custom machine learning code containing all of the calculations reported in the manuscript. The code is freely available at: <https://gitlab.com/yufeng.huang/sannp>. The training, validation, and testing data used in the manuscript are also included in the Data folder.

Detailed description about running the code is on the Gitlab website. Datasets used in this study have been provided in the example calculations.

S3. Convergence of the neural network errors in the size of the neural network

In the paper a large system with 550 input features and a 2-layer neural network with 500 nodes per layer is used. The reason for the large system is to avoid errors due to incompleteness. In the manuscript, the convergence with respect to the basis set size is shown

in Figure 4. Here we show how the forces converge when the number of nodes in the neural network increases. For this test, the $M=4$ piecewise cosine basis set, which corresponds to 56 input features, is used. The simplest case has 40 nodes in each layer in the 2-layer neural network. Changing the number of nodes only reduces the error by about 0.014eV/\AA .



S4. DFT data generation

The computational method used here is the same as in Kang and Wang [1].

The amorphous Si structure is first equilibrated in LAMMPS [2] using the ReaxFF reactive force-field [3], which allows for long simulations of amorphous structures with bond breaking and formation. We start with the $4 \times 2\sqrt{2} \times 2\sqrt{2}$ supercell of the crystalline diamond structure of silicon consisting of 256 atoms. Then the system is heated to 4300 K in 170 ps

with the NPT ensemble. After equilibrating the system for another 150 ps at 4300 K, the system is cooled down to 10 K in 1.1 ns. Amorphous Si is formed during the annealing process. Using the LAMMPS results at the initial structures, we perform DFT calculations in PWmat to relax the geometries. The results are consistent with experimental observations [4].

[1] J. Kang, L. W. Wang, *Physical Review B*, 96(2), 1–5. (2017)

[2] S. Plimpton, *J. Comput. Phys.* 117, 1 (1995).

[3] A. C. T. van Duin, S. Dasgupta, F. Lorant, and W. A. Goddard, *J. Phys. Chem. A* 105, 9396 (2001).

[4] S. Kugler, L. Pusztai, L. Rosta, P. Chieux, and R. Bellissent, *Phys. Rev. B* 48, 7685 (1993).