

Machine Learning for Atomic Simulation and Activity Prediction in Heterogeneous Catalysis: Current Status and Future

Sicong Ma and Zhi-Pan Liu*

Cite This: *ACS Catal.* 2020, 10, 13213–13226

Read Online

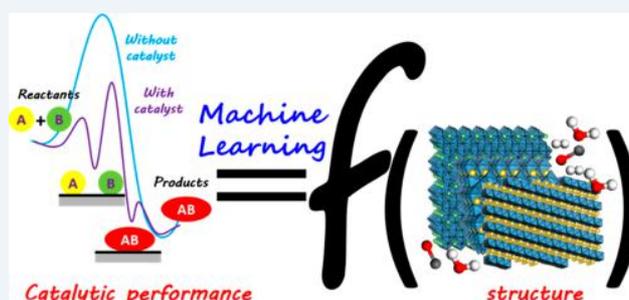
ACCESS |

Metrics & More

Article Recommendations

ABSTRACT: Heterogeneous catalysis, for its industrial importance and great complexity in structure, has long been the testing ground of new characterization techniques. Machine learning (ML) as a starring tool in data science brings new opportunities for chemists to interpret, simulate, and predict complex reactions in heterogeneous catalysis. Here we review the current status of ML methods and applications in heterogeneous catalysis by following two main streams: the top-down approach by learning experiment data and the bottom-up approach for making predictions from first-principles, which differ in the data source. We focus more on the latter, where ML interacts intimately with first-principles calculations for predicting the key properties (e.g., molecular adsorption energy) and evaluating potential energy surface (PES) to expedite the atomic simulation. The ML-based PES exploration represents the top gear that can largely replace the traditional roles of first-principles calculations for structure determination and activity evaluation but requires efficient methods for data set generation, sensitive structure descriptors to discriminate structures, and iterative self-learning to refine the ML potential. We illustrate these key ingredients of ML-based atomic simulation using the SSW-NN method developed by our group as the example. Three cases of SSW-NN application are presented to elaborate how ML can expedite the material and reaction simulation and lead to new findings on catalyst structure and reaction channels. The future directions of ML-based applications in heterogeneous catalysis are also discussed.

KEYWORDS: machine learning, heterogeneous catalysis, potential energy surface, density functional theory, global optimization, SSW-NN, LASP



1. CHALLENGES OF THEORETICAL CATALYSIS

Heterogeneous catalysis is where material science meets chemical reactions. It is no wonder that every progress in material characterization technique, from the earliest X-ray diffraction (XRD) to the latest transmission electron microscopy (TEM), brings great forward in catalysis science.¹ Theoretical simulations with density functional theory (DFT) calculations enter into the tool boxes of catalysis community in the past 30 years and have become an essential complement to experiment for understanding catalyst structures and catalytic activity. Not only are the reaction mechanisms clarified for a significant number of heterogeneous catalytic reactions,² but also general rules for predicting catalytic sites and reactivity are established.³ To date, much fundamental knowledge from well-defined, crystalline surfaces have been gleaned and the research focus shifts more toward catalytic systems under realistic conditions. In this regard, DFT-based theoretical simulation faces the ever-increasing challenges to cope with the rapidly increased degree-of-freedom caused by both the high structure complexity and the huge reaction space in catalysis.

The Haber–Bosch process, as a textbook example for illustrating the complexity of heterogeneous catalysis, utilizes alkali-promoted fused iron as catalyst for converting N_2/H_2 to ammonia.⁴ It generally contains multiple elements and multiple components: in addition to iron, the presence of alumina, calcium oxide, and potassium are also important to the performance; the catalysts are prepared from magnetite (Fe_3O_4) followed by high-temperature reduction. Apart from the complexity of catalysts, the reaction occurs at high temperature (500 °C) and high pressure (300 atm) with many intermediates, including adsorbed H and NH_x species. The past decades have seen many DFT-based calculations on Fe-catalyzed ammonia synthesis,^{5,6} which were however limited to single-crystal Fe surfaces (e.g., $Fe(111)$),⁷

Received: August 9, 2020

Revised: September 11, 2020

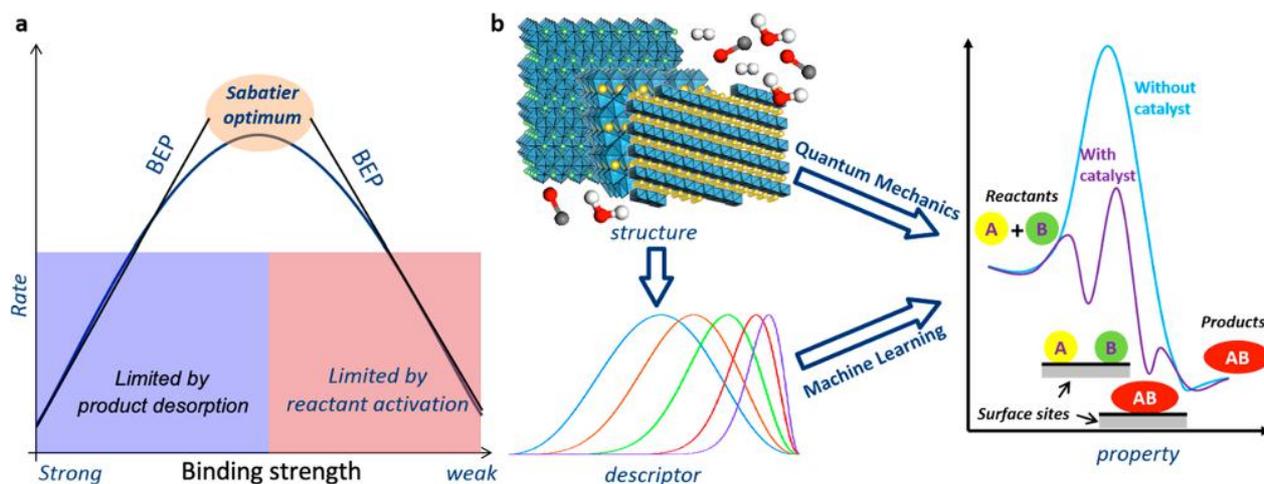


Figure 1. Schemes for BEP relation and Sabatier volcano plot (a) and machine learning for heterogeneous catalysis (b).

Fe(211)⁸) as what did in surface science experiment.⁹ Obviously, it would be highly desirable for theoretical simulations to predict the active catalyst structures under reaction conditions and to explore the intriguing questions relevant to catalysis, such as the roles of promoters and the effect of particle size.

In fact, the dilemma of catalytic reaction simulation is rooted in the fundamental kinetics laws governed by the Master Equation (eq 1).

$$\frac{d\mathbf{P}}{dt} = \mathbf{A}\mathbf{P} \quad (1)$$

where the rate of \mathbf{P} (a vector represents the population of different states) depends on the transition rate matrix \mathbf{A} . According to transition state theory (TST),^{10,11} the matrix element A_{ij} in transition rate matrix is the rate constant between i and j states, as summarized in eq 2.

$$k_{\text{TST}} = \frac{k_{\text{B}}T}{h} e^{-\Delta G_{\text{TS}}^{\ddagger}/k_{\text{B}}T} \quad (2)$$

It suggests that (i) the building of the matrix needs to run over all (important) states, the number of which explodes with the increase of system size and complexity; (ii) the occurrence of states separated by high barriers are exponentially slower compared to those with low barriers. For a reaction to occur at 500 °C with 1 s⁻¹ rate, the free-energy barrier is about 2 eV, which suggests that if the temperature drops to 400 °C, the rate is reduced to 0.01 s⁻¹. This answers the basic questions in catalysis (e.g., why ammonia synthesis requires high temperatures: the stepwise hydrogenations from N to NH₃ have an overall barrier of 1.8~2 eV).^{12,13}

In general, both the catalyst structure evolution under reaction conditions and the catalytic reactions of molecule belong to rare events with high reaction barriers, which require a long time-scale (from seconds to hours) in simulation in order to observe these processes. As *ab initio* (e.g., DFT) molecular dynamics (MD)¹⁴ relies on Born–Oppenheimer approximation to solve Schrödinger equations and Newtonian laws to move atoms, the typical time-scale is often limited to less than a nanosecond, i.e. 10⁶ times *ab initio* calculations with a typical time-step of 1 fs in MD. On the other hand, the computation speed of DFT, although is sensitive to the complexity of the employed density functional (e.g., PBE,¹⁵

HSE06¹⁶), has a poor scaling of at least $O(NnN)$ and is efficient only for small-sized systems within hundreds of atoms. Therefore, instead of straightforward MD, alternative low-cost approaches have to be taken in order to predict catalytic reactions from theory.

Since the late 1990s, the mainstream to study catalytic reactions converges to locate the transition state (TS) of reactions on potential energy surface (PES), from which the reaction barrier and the activity can be determined from theory.¹⁷ Many efficient TS search methods have been developed for heterogeneous catalytic reactions. They can be classified as the single-ended approach, i.e. starting from TS-like structure (the dimer method and its improved versions^{18–20}), and double-ended approach, i.e. starting from initial and final states (nudged elastic band^{21,22} and double-ended surface walking²³). As the TS search requires the information on the structure configurations, it becomes frustrated for reactions with complex reaction configurations such as those involved in surface structure restructuring and unusual reaction intermediates in reaction network.

Another popular way for activity prediction is based on the Bell–Evans–Polanyi (BEP) principle,²⁴ which avoids the time-consuming TS search by establishing the approximate linear correlation between the reaction barrier E_{a} and the reaction enthalpy ΔH (the energy difference between final and initial state), see eq 3.

$$E_{\text{a}} = a\Delta H + b \quad (3)$$

Between 2001 to 2003, several groups (e.g., Neurock, Hu, and Nørskov groups, etc.)^{3,25–27} determined independently the parameters of a and b in BEP relation by DFT for a large number of surface reactions, which provides the quantitative basis for plotting the Volcano curve of activity, the Sabatier principle known since 1910s, as shown schematically in Figure 1a. The binding strengths of products and reactants need to be in a fine balance, neither too strong nor too weak, to achieve the highest activity.

The BEP approach has been further extended to predict the activity of complex multistep reactions by using simple descriptors that can be computed facily. For example, the d-band center (ϵ_{d}) proposed by Hammer and Nørskov²⁸ was known to be a good indicator for the molecular adsorption

strength on metal surfaces, which can be computed by averaging the occupied d-states (n_d) in eq 4.

$$\varepsilon_d = \frac{\int_{-\infty}^{\varepsilon_f} n_d(\varepsilon) \varepsilon d\varepsilon}{\int_{-\infty}^{\varepsilon_f} n_d(\varepsilon) d\varepsilon} \quad (4)$$

Except for the electronic descriptors, the geometry-based descriptors are of most interests. The generalized coordination number (\overline{CN}) proposed by Sautet group^{29,30} connects the atomic coordination environment (eq 5) with the binding energies of reaction intermediates for oxygen reduction reaction.

$$\overline{CN} = \sum_{j=1}^n \frac{cn(j)}{cn_{max}} \quad (5)$$

The sum includes all of the first-nearest neighbor atoms ($cn(j)$) and the division by the maximum number of first-nearest neighbors in the bulk (cn_{max}) ensures that \overline{CN} spans the range between 0 and 12 in face-centered cubic metals. Based on these simple descriptors, the high-throughput screening techniques were then developed for fast evaluating the activity of different materials across the periodic table.^{31,32}

After obtaining the elementary reaction data from first-principles (e.g., the reaction energy and barrier), the mean-field microkinetics or kinetic Monte Carlo (KMC) can be utilized to simulate chemical reactions on surfaces, which will yield the time-resolved catalytic activity and selectivity and provide the fine details on the surface status during reaction dynamics. Although the first-principles data-based KMC simulations remain largely limited to few well-defined crystalline surfaces and relatively simple reactions, the success of these works by carefully benchmarking with experimental kinetics data do represent a marvelous hallmark in theoretical catalysis.^{33,34} Apart from the difficulty in considering the structural complexity of the catalysts and the tedious computation of the kinetics data for all elementary steps, a major concern for kinetics simulations originates from the intrinsic error of DFT calculations (~ 0.1 eV error) in computing adsorption energies, reaction barriers, and vibrational entropies for certain molecules/reactions (e.g., CO adsorption). These DFT errors have to be corrected empirically in kinetics simulations in order to yield a more realistic rate for the experimental conditions.

With the development of method and the massive production of data, the 21st century has witnessed a great advance in data-driven science (e.g., machine learning (ML)) and its applications in catalysis research. Similar to the high-throughput approaches that use geometry descriptors for finding trends, ML methods also utilize the geometry-based information for the energy and the property prediction, which renders the high speed in computation. ML applications on energy materials,³⁵ catalysis informatics,³⁶ and heterogeneous catalyst design and discovery^{37,38} have been reviewed recently. Different from these previous contributions, here we will focus more on the progress of theoretical catalysis via atomic simulation, which is now benefited significantly by ML techniques. We will review ML-based atomic simulation methods utilized in catalysis, in particular stochastic surface walking (SSW) global optimization based on neural network (NN) potential method (SSW-NN) developed in our group.^{39,40} These ML methods announce the advent of a post-DFT era in theoretical catalysis, which aim to solve

challenging catalysis problems that are traditionally intractable from first-principles calculations, either in the time-scale or in the size and the structural complexity of the system.

2. ML FOR HETEROGENEOUS CATALYSIS

As an interdisciplinary technology, ML was invented in the fields of computer science and statistics science,⁴¹ which aims to identify trends, rules, or functions from the existing data, instead of from physical laws. With dozens of ML algorithms developed to date, from kernel ridge regression to support vector regression and to sophisticated deep NN, ML techniques have been successfully applied to many fields of science that routinely have access to big data.

The applications in heterogeneous catalysis started in 1990s,^{42,43} mainly on learning experimental data. For example, Kito et al.⁴⁴ used NN to predict the product distribution of ethylbenzene oxidative hydrogenation based on experimental catalyst data. They chose different products, including styrene, benzaldehyde, benzene, toluene, CO, and CO₂ as the outputs of NN, and nine different independent variables as input, including the valence, surface area of the catalyst, amount of catalyst, typical valence, ionic radius, coordination number, electronegativity, partial charge of oxygen ion, and standard heat of formation of oxides. The high cost in obtaining a large, consistent data set is the major bottleneck for ML applications in experiment. Compared with experiment, the data produced from DFT calculations is self-consistent and much more economic to obtain. This is particularly advantageous for materials and reactions that are unlikely to synthesize or to occur under ambient conditions. Furthermore, many material gene databases with DFT results are being established around the world,⁴⁵ which, as a speed-up, enables the convenient data mining for ML of novel materials.

Broadly speaking, there are two branches for ML applications in catalysis (i.e., the top-down and bottom-up approaches). The former is to directly predict catalytic performance based on existing experimental data as those started from 1990s. The latter learns the results from quantum mechanics (QM) calculations, aiming to predict property and evaluate PES, which can be regarded as a replacement of QM for complex catalysis applications, as shown in Figure 1b.

2.1. Top-Down Approach: Learning from Experimental Data. To build a ML model for catalytic performance prediction, it is essential to first define the data descriptors for catalyst. Traditional concepts for catalyst include, for example, the particle size, surface area, valence state, surface composition and coordination, which can be determined from advanced characterization techniques, such as XRD, TEM, Brunauer–Emmett–Teller surface area analysis, X-ray photoelectron spectroscopy, and so on. However, because of the high cost of structure characterization, it is a common practice that only the catalysts with good catalytic performance are characterized in detail, causing the lack of data for most catalysts reported in literatures. In practice, the low-cost descriptors are adopted in ML applications, typically from the catalyst synthesis conditions (e.g., feed composition and ratio, synthesis time and temperature, pH value, etc.) and the operation conditions of reaction (e.g., catalyst amount, reaction temperature, time, reactant amount, etc.).

For instance, Sasaki et al.⁴² showed that the yield and byproducts of NO decomposition over the Cu/ZSM-5 zeolite catalyst can be predicted by a NN model using the experimental conditions as input such as compositional

quantity, temperature, and pressure. Similarly, Rahman et al.⁴⁶ used the temperature, reaction time, substrate molar ratio, and enzyme amount as the inputs for a NN model, which can predict the yield of lipase-catalyzed diacylglycerol synthesis. These models help to optimize the synthesis conditions and reaction conditions to maximize the reaction activity. However, they are not expected to be transferable (i.e., only work for the experimental data for each research group) because the consistency of data cannot be guaranteed due to some hidden variables (e.g., stirring rate, experimental equipment).

ML-assisted literature analysis (i.e., the data mining technology) is a possible solution to overcome the data limitation and eliminate the inconsistency between data. By using the text mining and information extraction of natural language processing, Kim et al.⁴⁷ established a ML model using natural language processing and decision tree to predict the critical parameters for synthesizing oxide materials from 12 000 scientific articles about metal oxide synthesis (Figure 2). For

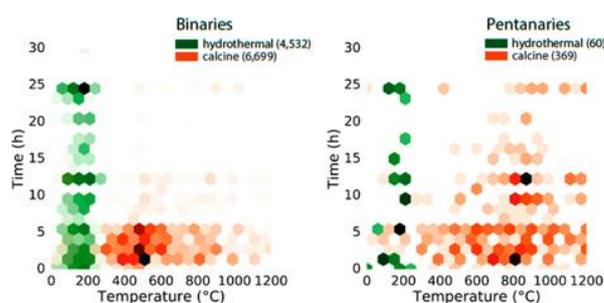


Figure 2. Two-dimensional hexagonally binned normalized histograms of hydrothermal reaction and calcination times and temperatures for binary and pentanary oxides. Figure adapted from ref 47. Copyright 2017 American Chemical Society.

instance, they found that most hydrothermal reactions are carried out between 150 and 200 °C for 12 or 24 h. The hydrothermal condition used to synthesize both simple and complex oxides occur at the similar and only modestly high temperatures but often with fairly long times, and the calcination temperature is typically material-specific and driven by the structural change. Yildirim and co-workers⁴⁸ collected 5508 experimental data for steam reforming of CH₄ from literatures. The database was then analyzed using decision trees to extract correlations and trends, where 21 variables related to catalyst preparation and operational conditions were used as input variables. Similar applications have also been reported in other reactions (e.g., CO oxidation, CO₂ reduction, etc.).^{49–51}

It should be emphasized that the conclusions from the top-down ML approaches rely on the scope of published experimental data: it is virtually impossible to predict the performance of unknown catalysts and find unknown reactions. The emergence of ML robots⁵² might be the solution for collecting a large amount of experimental data without human input. There is much to explore in the future of automated experiments.

2.2. Bottom-Up Approach: Predicting from First-Principles. It is more attractive to make prediction from first-principles, i.e. without input from experiment. Such a standard ML procedure would consist of four steps: (i) building a data set from QM calculations, often DFT, which should contain various structures and their properties (e.g., energy, band gap, d-band center, etc.); (ii) identifying the

features that can link the structure with the properties for prediction. These features can be basic constant of atom/elements, the geometrical coordinate or other low-cost computable properties (e.g., atomic charge); (iii) designing ML models and training ML models on the data set; (iv) predicting properties using ML models for unknown structures.

Typically, step (i) is the most time-consuming step, which requires heavy QM computation on many candidate structures. We will show later that the efficient PES sampling techniques are critical for obtaining a desirable data set. The steps of (ii) and (iii) are application-orientated where the chemical intuition, the experience, and also trial-and-error tests are typical recipes to identify the suitable features and to construct appropriate ML models. In the following, we will overview the ML literatures on heterogeneous catalysis from two aspects according to the output of ML models (i.e., adsorption energy prediction and PES prediction).

2.2.1. Adsorption Energy Prediction. Based on the BEP principle, i.e. the linear dependence of the reaction barrier on the adsorption energy, the adsorption energy is regarded as a key parameter in heterogeneous catalysis and is also a low-cost computable property. The successful prediction of adsorption energy can lead to a quick evaluation of catalytic activity of system. The numerical accuracy for the adsorption energy prediction should be high (e.g., below 0.1 eV) because the value itself is typically -3 to 1 eV and an error of 0.1 eV may totally change the activity.

For finding the active site of CO reduction on NiGa alloy, Ulissi et al.⁵³ developed a NN-aided discovery approach which relates the CO adsorption energies with different NiGa bimetallic surface sites. The surface site is described by seven structural descriptors, i.e. the coordination number (CN) of Ni, CN_{Ga}, the average CN for Ni–Ni bonds (CN_{Ni–Ni}), the average CN_{Ni–Ga}, the average CN_{Ga–Ni}, the average CN_{Ga–Ga} and the fraction of the Ni in the alloy. This led to the discovery of a previously unconsidered active site, single Ni atoms surrounded by surface Ga atoms, which is shown to exhibit a step-like kinetic behavior and the best thermodynamics for CO reduction.⁵⁴ Jinnouchi and Asahi^{55,56} predicted the rates of NO decomposition as a function of composition and particle size by using a ML scheme, where the combination of Bayesian linear regression and the local similarity kernel is utilized to connect the structure with the O, N, and NO adsorption energies. The catalytic activities of NO decomposition on Rh_{1–x}Au_x alloy nanoparticles were then estimated by Sabatier analysis based on the predicted adsorption energy and the available BEP relationship.

Apart from the catalytic activity estimation on a single catalyst, the ML-based adsorption energy prediction can be extended for the purpose of the high-throughput catalyst screening and the reaction network optimization. For catalyst screening, Tran and Ulissi⁵⁷ constructed the relationship between structural fingerprints and the CO and H adsorption energies. Guided by the optimal adsorption energy of CO and H ($E_{\text{CO}} = -0.67$ eV and $E_{\text{H}} = -0.27$ eV) for CO₂RR and hydrogen evolution reaction (HER), respectively, they performed a systematic screening of alloys with 31 different elements that encompasses 50% of *d*-block elements and 33% of *p*-block elements. Finally, 131 candidate surfaces from 54 bulk alloys for CO₂ reduction and 258 surfaces from 102 bulk alloys for H₂ evolution were identified. This method successfully accelerated the design of the CuAl alloy catalyst

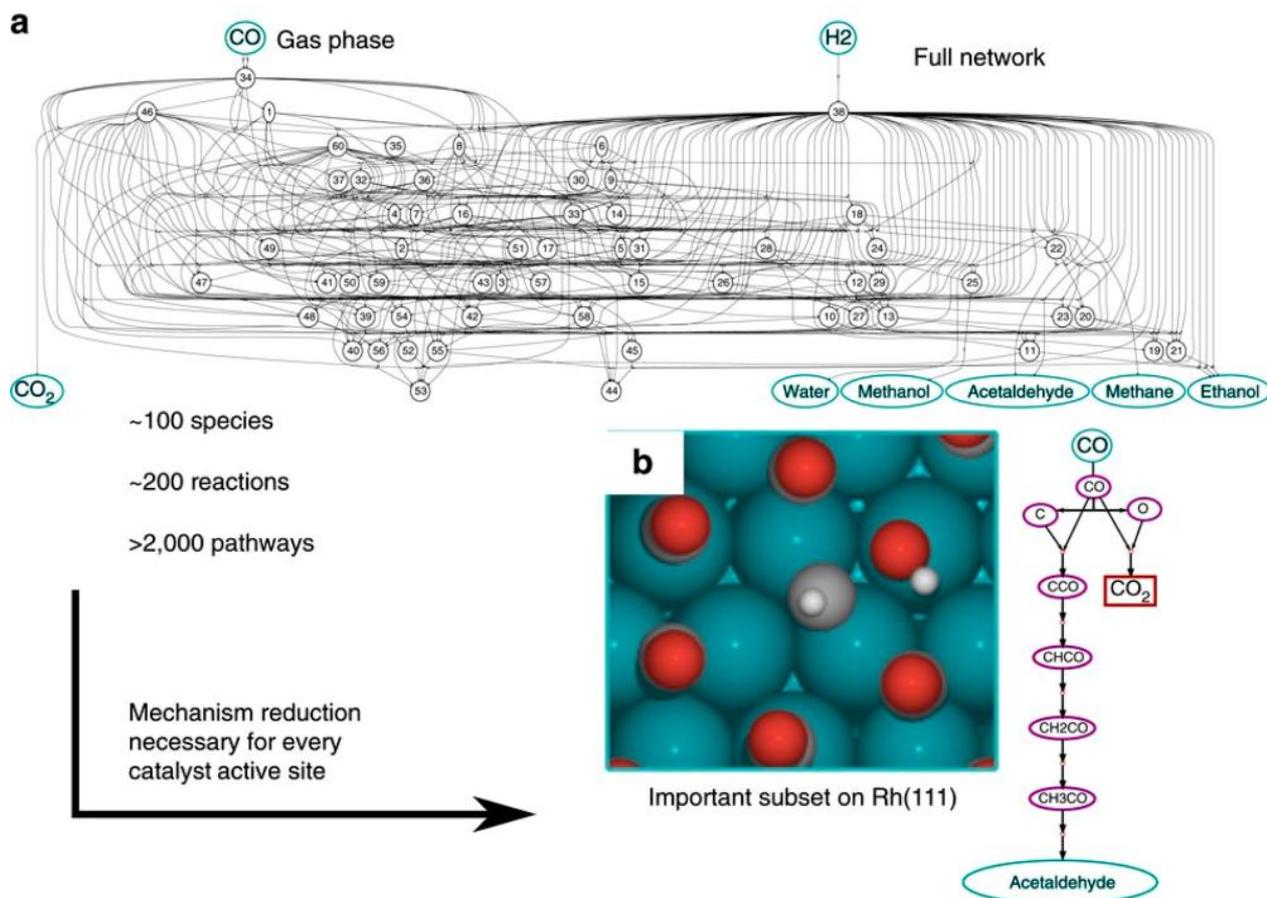


Figure 3. (a) Reaction network for the reaction of syngas ($\text{CO} + \text{H}_2$) to CO_2 , water, methane, methanol, acetaldehyde and ethanol, including surface-adsorbed intermediates with up two carbons and two oxygens (C1/C2 chemistries). (b) The reduced network for syngas reactivity on Rh (1 1 1), producing acetaldehyde selectively as confirmed by the experiment. The reduction of the reaction network (a) to the reduced reaction network (b) is achieved using a machine learning aided reaction network optimization framework. Figure adapted from ref 60. Copyright 2017 Nature Publishing Group.

which provides multiple sites and surface orientations with near-optimal CO binding for both efficient and selective CO₂ reduction.^{58,59} For finding reaction mechanism, Nørskov and co-workers⁶⁰ proposed a reaction network optimization framework using Gaussian process regression (GPR) to study the reaction of syngas ($\text{CO} + \text{H}_2$) over Rh (111) catalysts under experimentally relevant operating conditions (573 K and 1 atm of gas phase reactants), as shown in Figure 3. Starting from a few DFT energies of the intermediates in the reaction network, a computationally inexpensive GPR scheme was used to predict the free energy for all intermediates in the reaction network. The BEP relation was exploited to estimate the activation energies for all reactions in the network, and a simple classifier was used to select the potential rate-limiting steps (Figure 3a). Through an iterative GPR model refinement process, where only potential rate-limiting steps were analyzed via further DFT calculations, a probable reaction network from syngas to acetaldehyde was finally identified (Figure 3b).

Instead of the quantitative prediction of adsorption energy, ML using the same data set of adsorption can be utilized to provide qualitative understanding on what controls the adsorption strength. Toyao et al.⁶¹ developed a simple ML model with 12 descriptors for predicting the adsorption energies of CH_x and H species on Cu-based alloys. Among 12 descriptors, the element group, surface energy, and melting

point were found to be the key factors affecting the adsorption energy (Figure 4a). Rappe and co-workers constructed a regularized random forest ML model to quantify the importance of different descriptors of the Ni₂P (0001) surface structure on determining the free energy of H adsorption. The results showed that the Ni–Ni bond length is the most important descriptor for HER activity; a shorter Ni–Ni bonds give higher HER activity (Figure 4b).⁶² Similarly, Xin et al.^{63,64} built a ML model with a large number of different descriptors for predicting CO binding on alloys. They identified the important role from the d-band shape and sp-band filling on CO binding, even for coinage metals. Indeed, this information provides the alternative to the adsorption energy for activity prediction.

Since the adsorption energy can be related to d-band center for metals, ML is also utilized to predict directly the d-band center as to provide a more general model for different adsorbates. Takigawa et al.^{65,66} developed a ML-based model by gradient boosting regression method to predict the d-band centers for 11 metals (Fe, Co, Ni, Cu, Ru, Rh, Pd, Ag, Ir, Pt, Au) and their pairwise bimetallics for two different structures (metal impurities and overlayer-covered metal surfaces). It is shown that the d-band centers are reasonably well predicted with average root-mean-square error (RMSE) less than 0.5 eV, even only six descriptors and 25% data given for training. In

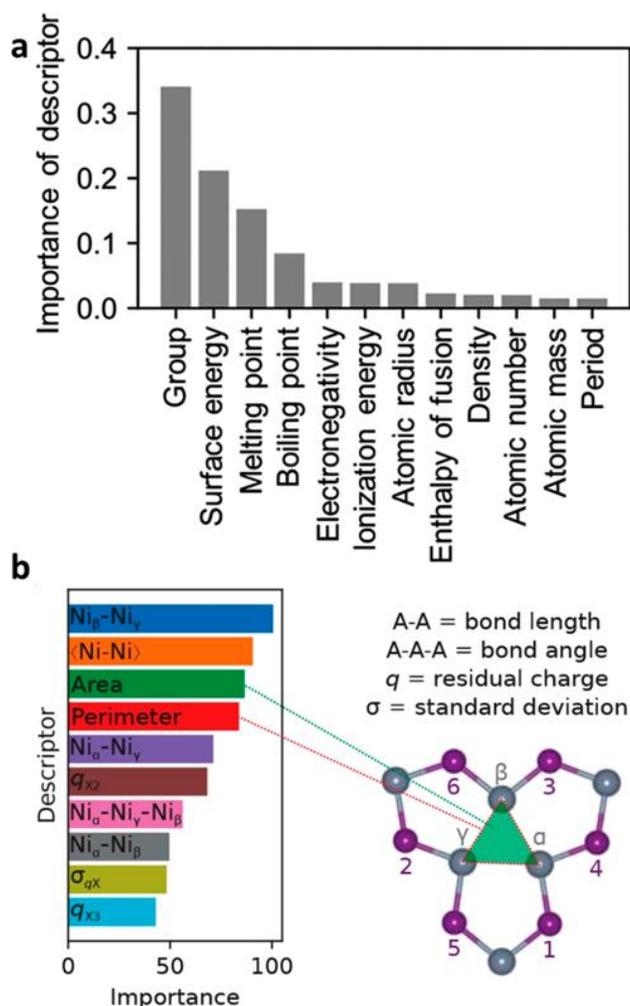


Figure 4. (a) Feature-importance scores of the descriptors for the extra tree regression prediction of the adsorption energies of CH₃ on Cu-based alloys. (b) Relative importance of descriptors calculated from RRF model. Only the top 10 features are shown. Figure adapted from ref 61 (a) and ref 62 (b). Copyright 2018 American Chemical Society and Copyright 2018 American Chemical Society.

the same idea, Sun et al.⁶⁷ attempted to identify the better spinel oxide catalysts for oxygen evolution reaction (OER). They first observed that the activity on spinel oxides is related to the covalency competition between tetrahedral and octahedral sites. This competition led to an asymmetric M_T-O-M_O backbone where the weaker covalent M-O bond in lattice facilitates the exposure of the cation site and therefore enhances the OER activity. The covalency strength can be quantified by the energy difference between the centers of the metal d and oxygen p bands, denoted D_M . On the basis of a set of structure features of spinel, a ML model was thus constructed to predict the D_M with a mean absolute error of 0.05 eV. They thus predicted the [Mn]_T[Al_{0.5}Mn_{1.5}]O₄ spinel catalyst to have an optimal D_M , and the subsequent experimental results confirmed its superior activity.

Overall, the activity prediction using the adsorption energy and the alternatives is an approximated approach built on the BEP relation. Considering the exponential dependence of the reaction rate on reaction barriers, the other crucial factors, such as the catalyst stability, the active site structure, and the exact reaction kinetics, cannot be totally neglected, which indeed

become increasingly important for evaluating the peak activity of volcano curve. The atomic simulation remains as the key tool to go beyond BEP relation in order to determine the activity under realistic conditions.

2.2.2. Atomic Simulation for Fast PES Exploration. QM and statistic mechanics are two branches of fundamental sciences that play on the battle ground of PES. The former provides the energy and latter derives thermodynamics from the energy landscape. It is of central significance to develop fast and reliable PES evaluation methods to meet the ever-increasing demand for simulating complex systems, such as those in heterogeneous catalysis. ML-based atomic simulation emerged in recent years is such a promising field, which relies on a ML potential for evaluating PES, an analytic numerical function with numerous adjustable parameters. Compared with empirical force fields, the ML potential is advantageous in studying complex material systems as well as chemical reactions as long as the training data set contains the related PES data.

The energy expression to correlate structure with energy is basic to the ML model for PES evaluation. Since energy is an extensive quantity, the analytic expression of energy can be conventionally expanded as the summation of different functional forms constituted by local geometrical variables, such as bond, angles, and dihedral angles. The many-body expansion (MBE) method is a typical approach, which, being identical to the force field methods, splits the total energy as the interaction terms of one-, two-, and three-body terms (E_i , E_{ij} , and E_{ijk} , i, j, k are indices of atom), respectively, as shown in eq 6.

$$E = \sum_i E_i + \sum_i \sum_{j>i} E_{ij} + \sum_i \sum_{j>i} \sum_{k>j} E_{ijk} + \quad (6)$$

Instead of using physically meaningful functions, each term utilizes complex numerical functions, the parameters of which can be trained via ML techniques. For example, Bartok et al. used Bayesian inference techniques to correct DFT one-body and two-body energies for water, which significantly improves the description of the structure and dynamics of liquid water.⁶⁸ Because of the sharp increase of the many-body terms (four-body and above) with the increase of system size, the MBE method is problematic for describing large size systems and also those with salient many-body bonding characteristics (metallic systems) and thus not suitable for heterogeneous catalysis.

Behler and Parrinello^{69–72} proposed a high-dimensional neural network (HDNN) scheme in 2007, as shown in Figure 5, that splits the total energy as the summation of individual atomic energy, as shown in eq 7.

$$E = \sum_i E_i \quad (7)$$

where E_i for each atom is the output of a standard feed-forward NN. The input of NN is a set of structural descriptors to describe the atom bonding environment. The NN parameters can be trained using the first-principles PES data set. Compared with the MBE method, the atomic energy summation method is linear scaling with respect to system size and thus is suitable for simulating large material systems. In HDNN architecture, the functional form of NN can be replaced by other numerical representations, especially when target systems are relatively simple. For example, GPR proposed by Bartok et al.^{29,38,39} have been successfully utilized

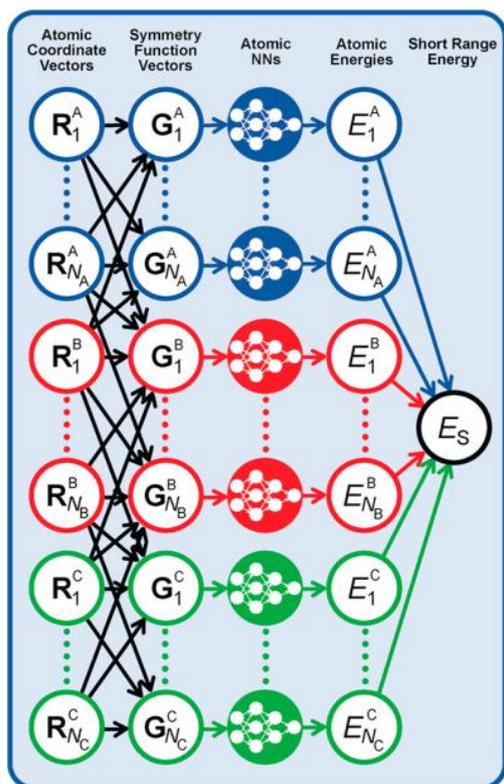


Figure 5. Schematic structure of HDNN architecture for a ternary system. Figure adapted from ref 72. Copyright 2017 Wiley Publishing Group.

for a number crystalline systems, functional materials, and amorphous solids.^{73–75}

To describe the local atomic environment, Behler proposed a series of Gaussian-type symmetry functions as the input of NN. The most used two-body G^2 and three-body G^3 are shown in eqs 8–10,⁶⁹ where r_{ij} is the internuclear distance between atoms i and j and θ_{ijk} is the angle centered at the i atom with neighbors j and k . These structural descriptors are invariant with respect to the permutational, rotational, and translational movement of the system and thus can produce a continuous PES for deriving the first energy derivatives (i.e., the atomic force):

$$f_c(r_{ij}) = \begin{cases} 0.5 \times \tanh \left[1 - \frac{r_{ij}}{r_c} \right], & \text{for } r_{ij} \leq r_c \\ 0 & \text{for } r_{ij} > r_c \end{cases} \quad (8)$$

$$G_i^2 = \sum_{j \neq i} e^{-\eta(r-r_c)^2} \cdot f_c(r_{ij}) \quad (9)$$

$$G_i^3 = 2^{1-\zeta} \sum_{j,k \neq i}^{all} (1 + \lambda \cos \theta_{ijk})^\zeta \cdot e^{-\eta(r_{ij}^2 + r_{ik}^2 + r_{jk}^2)} \cdot f_c(r_{ij}) \cdot f_c(r_{ik}) \cdot f_c(r_{jk}) \quad (10)$$

With the similar idea, other local geometry representations (e.g., bispectrum and smooth overlap of atomic positions)⁷³ were also proposed for PES construction.

While HDNN was initially developed for material systems, it has been utilized for catalysis-related problems. Bose and Kitchin⁷⁶ investigated the oxygen interaction with Pd (111) by

constructing a PdO HDNN potential. The grand-canonical MC simulations on a (10 × 10) Pd slabs at 600 K was performed, which found that the mean oxygen coverage is around 0.50 monolayer. The oxygen diffusion barrier from HDNN potential is 0.40 eV, which are in agreement with experimental results of 0.4–0.5 eV.⁷⁷ Artrith et al.⁷⁸ utilized HDNN potential for optimizing the structure of Cu/ZnO. The canonical (NVT) MD simulation at 1000 K was carried out for a large Cu cluster (612 atoms) supported on ZnO (10 $\bar{1}0$) surface, a system of 7524 atoms. The yielded Cu cluster is highly distorted, particularly at the interface, which provides a structural model for understanding the true catalyst morphology under reaction conditions.

Undoubtedly, a major problem for ML-based atomic simulation lies at the high cost of producing the large PES data set required for training ML potential, which must involve extensive QM calculations for a large number of structural configurations. Ideally, the PES data set for ML should be representative and compact but needs to be as global as possible to cover both minima and transition regions on PES. In 2017, our group proposed to utilize the SSW^{79,80} global optimization trajectories for constructing the representative PES data set, and the combination with NN technique (SSW-NN) for atomic simulation turns out to be successful for simulating a wide range of materials and reactions and also for guiding experiment.^{81,82} The SSW-NN method is now implemented in LASP software (Large-scale Atomic Simulation with neural network Potential, accessible from www.lasphub.com), which incorporates the PES data generation, NN potential training, and the atomic simulation in one platform. By using LASP, a large set of global NN (G-NN) potentials is now established via the automated global-to-global learning procedure. The G-NN potentials cover a wide range of elements across the periodic table. In the following sections, we will introduce the SSW-NN method and its three applications to heterogeneous catalysis.

3. SSW-NN FOR HETEROGENEOUS CATALYSIS

3.1. SSW-NN Methodology. In principle, the global PES sampling methods would be an outstanding choice for PES data generation. Simulated annealing has been the most popular method, which explores PES via repeated heating and cooling cycles.⁸³ The enhanced MD method such as metadynamics^{84,85} can be the valuable supplement to add the reaction data. Nevertheless, MD sampling for data generation still suffers from the “short-sighted” problem because of the exponentially low probability to overcome high reaction barrier at low temperatures and the preference of trapping at high-entropy structure regions at high temperatures. As a result, the PES data thus generated are often overwhelmingly redundant, being highly localized to a few initial phases. This will inevitably lead to the inadequacy of thus-obtained ML potentials for predicting unknown materials and reactions. However, the other global optimization methods, such as basin-hopping,⁸⁶ evolutionary algorithm,⁸⁷ genetic algorithm,^{88,89} and the particle swarm optimization method (CALYPSO),⁹⁰ have also been tested for on-the-fly ML model training of materials in recent years. These methods transform the PES by overlooking the transition region between minima to realize fast global minimum (GM) search. Generally, they sample the PES data from the structural relaxation trajectories and thus may well miss key reaction channels.

Our group developed SSW method in 2013 for global optimization of aperiodic systems,⁸⁰ such as molecules and clusters, and the method has been extended to periodic crystals.⁹¹ Compared with other unbiased global optimization methods, SSW algorithm⁷⁹ can explore unbiasedly both minima and saddle points on PES due to the small step-size utilized in structure displacement. In SSW, an automated climbing mechanism is implemented to manipulate a structural configuration moving smoothly from a local minimum to a high-energy configuration along one random mode direction.

The climbing mechanism of SSW method learns from the TS location method, the bias-potential driven constrained-Broyden-dimer (BP-CBD) method,²⁰ that utilize bias potentials to overcome the barrier between minima. In one particular SSW step, a modified PES, as shown in Figure 6a, is

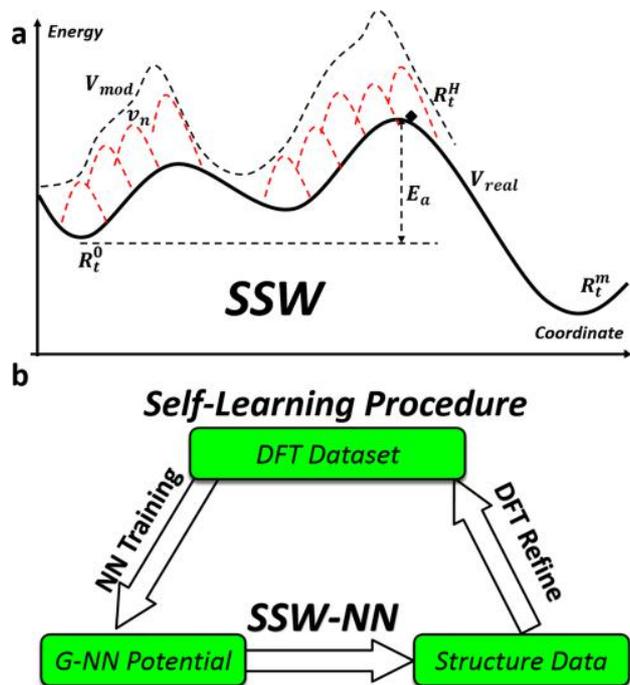


Figure 6. (a) Illustration of the SSW global optimization on 1-D PES. (b) Self-learning procedure of SSW-NN method for obtaining G-NN potential.

utilized for moving from the current minimum, R_t^0 , to a high energy configuration, R_t^H (the climbing), in which a series of bias Gaussian potential is added one by one consecutively along the direction. Once the R_t^H is reached, all bias potential are removed and the local optimization is performed to quench the structure to a new minimum. Different from the BP-CBD method, each SSW step (from one minimum to another) will choose a random direction to perturb the structure after the direction is refined (softened) using the biased-CBD method. At the end of the SSW step, a structure selection module (e.g., in Metropolis Monte Carlo scheme) is applied to accept/refuse the new minimum.

Inherited from SSW global optimization, the SSW-NN method is an iterative global-to-global approach for generating G-NN potential and performing the global optimization (Figure 6b). The G-NN potential follows the HDNN framework, as introduced in Figure 5, but the input layer utilizes a more sophisticated power-type structure descriptors

(PTSD),⁸¹ which is developed to be compatible with the SSW global optimization data set. In PTSD, not only are the traditional two-body and three-body terms included but also the four-body terms are added and the spherical functions are introduced to enhance the structure discrimination. For example, eqs 11–13 show the two-body terms utilized in PTSD, where R^n is the power function to describe the radial distribution and the Y_{Lm} is the spherical function to take into account the angular distribution of the chemical environment.

$$R^n(r_{ij}) = r_{ij}^n f_c(r_{ij}) \quad (11)$$

$$S_i^1 = \sum_{j \neq i} R^n(r_{ij}) \quad (12)$$

$$S_i^2 = \left[\sum_{m=-L}^L \left| \sum_{j \neq i} R^n(r_{ij}) Y_{Lm}(r_{ij}) \right|^2 \right]^{1/2} \quad (13)$$

In the beginning of SSW-NN self-learning procedure (Figure 6b), the NN potential is initially obtained by learning a small data set (typically less than 1000 structures) collected from short-time SSW sampling based on DFT, which are often restricted to small systems (below 20 atoms) of known configurations. The data set needs to be calculated by DFT with a high-accuracy setup. Next, the SSW global optimization based on NN potential will be carried out extensively, starting from a variety of initial structures, mainly randomly constructed, with different morphology, including bulk, surface and clusters, different chemical compositions, and different number of atoms per cell. After each iteration of global optimization, a small data set with diverse structures on PES is screened out by selecting either randomly or from those exhibiting new atomic environment (e.g., out-of-bounds in structural descriptors, unrealistic energy/force/curvature). These additional data will be calculated by DFT with the same high-accuracy setup and then added to the training data set for a new iteration of NN potential update. Typically, more than ~100 iterations are required to finally obtain a transferable G-NN potential. The accuracy for G-NN potential is typically 5~10 meV/atom for RMSE of energy and 0.1~0.2 eV/Å for RMSE of force.

3.2. LASP Software. LASP⁴⁰ software project initiated in the early of 2018. The initial objective is to merge two major simulation tools, SSW global optimization and G-NN calculations, developed by our group into a single package for the better and simpler usage. LASP is now shaped toward a software platform for many purposes, not only for atomistic simulation but also for PES data building and exchange and even for G-NN potential generation. A large set of powerful simulation techniques has been assembled into LASP program (the current version 2.2) to simplify the usage and to enlarge the scope of the current atomistic simulation. This can be seen in Figure 7, which overviews the current modules in LASP. Apart from the G-NN computation for energy and force evaluation, LASP implements standard data-exchange interfaces to connect with common PES evaluation packages, which allows for the PES data generation using QM calculation. For PES exploration at the finite temperature and pressure conditions, LASP also provides the standard and restrained MD functionalities, in addition to the structure local/global geometry optimization and TS search. All these functionalities can be switched on and off in LASP with short keywords, and

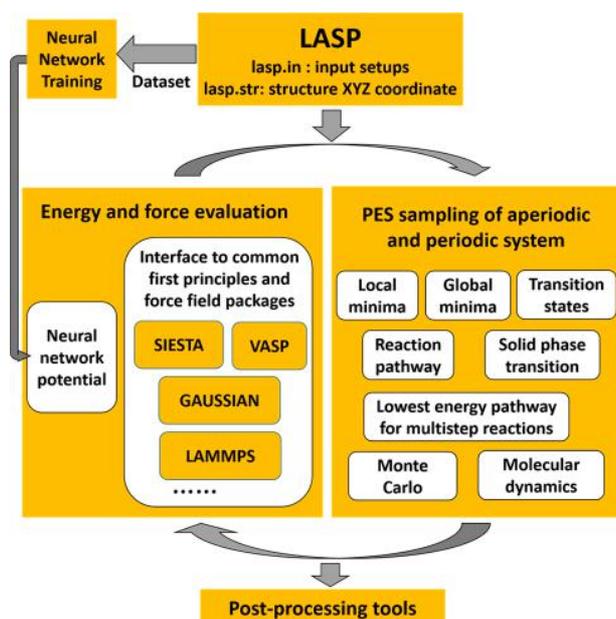


Figure 7. Architecture of LASP code and the modules implemented.

thus a variety of different tasks can be fulfilled flexibly within one platform, such as SSW-NN PES global exploration, PES data generation using QM calculations, G-NN potential creation, and MD simulation based on G-NN potential.

3.3. SSW-NN Applications. 3.3.1. Case 1: Y-Stabilized ZrO_2 Structural Evolution. Y-stabilized ZrO_2 (YSZ) is one of the most important ion conducting solids for solid oxide fuel cells and solid oxide electrolysis cells.^{92–94} The composition and temperature-dependence of the oxygen anion transportation have been the focus in a large volume of literatures.^{95–97} The composition and temperature-dependence of the oxygen anion transportation have been the focus in a large volume of literatures. The observations from experiments are, however, at odds with the computational results from empirical potentials, where the computed oxygen diffusion coefficient, reaches the maximum generally below the maximum observed in experiment (i.e., 8 mol % Y_2O_3 (8YSZ) below 1700 K). To resolve this puzzle, we recently established the ternary Y–Zr–O G-NN potential based on the SSW global PES data set that covers a wide range of $Y_xZr_yO_z$ compositions with different structural types (e.g., bulks, surfaces, clusters), including Zr , ZrO_x , Y_2O_3 and $Y_{2x}Zr_{1-2x}O_{2-x}$ mixed oxides with $Y/Zr \sim 1:9$ to $\sim 4:3$. The final training data set consists of 28 803 structures. The G-NN has a five-layer network (188-60-50-50-1) for each element, reaching to 71 103 fitting parameters in total. For the final G-NN potential, the RMSE for the energy and the force reaches 7.674 meV/atom and 0.165 eV/Å, respectively.^{98,99}

SSW-NN global search was then utilized to clarify the thermodynamics hull diagram for bulk $Y_{2x}Zr_{1-2x}O_{2-x}$ as shown in Figure 8a, where the relative formation energy of $Y_{2x}Zr_{1-2x}O_{2-x}$ is plotted against $Y_2O_3\% = Y_2O_3/(Y_2O_3+ZrO_2)$.⁹⁸ It shows that 8YSZ, the composition to have the highest ion conductivity in experiment, is in fact not the thermodynamically stable phase, which prefers to disproportionate to the neighboring compositions, 6.7YSZ and 20YSZ. This explains the long-term instability of 8YSZ as observed in experiment. By further performing long-time MD simulation, we have computed the anion diffusion coefficient

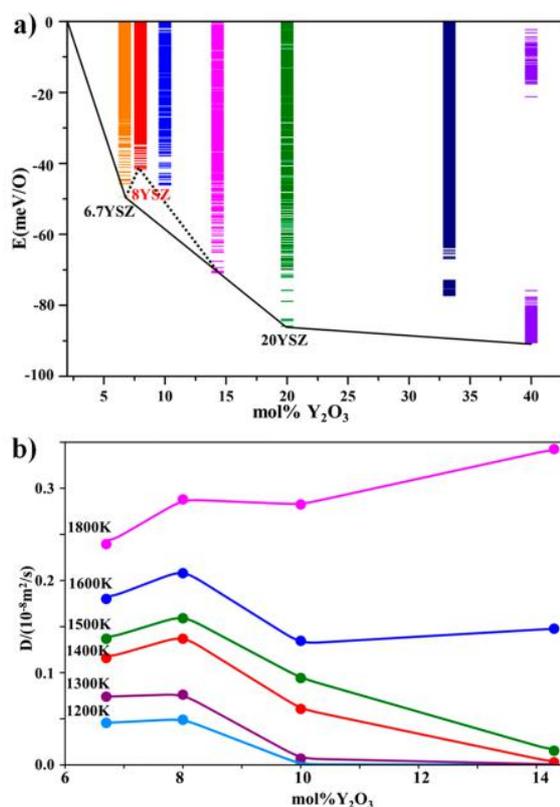


Figure 8. (a) Thermodynamic convex hull diagram for $Y_{2x}Zr_{1-2x}O_{2-x}$ with respect to the cubic- ZrO_2 and cubic- Y_2O_3 . (b) Oxygen diffusion coefficient D vs the concentration of Y_2O_3 at different temperatures. Figure adapted from refs 98 (a) and 99 (b). Copyright 2020 American Institute of Physics and Copyright 2020 American Chemical Society.

for different YSZ samples at up to 8 different temperatures with a long time-scale up to 10 ns. The calculated diffusion coefficient D is plotted against the concentration of Y in Figure 8b from 1200 to 1800 K.⁹⁹ The G-NN simulation shows that the D of oxygen anion peaks correctly at 8YSZ for a wide temperature range below 1600 K, reproducing the key finding in experiments.⁹⁷ With the increase of temperature to 2000 K, the maximum of D curve shifts from 8 mol % to the higher dopant concentrations, also in agreement with experimental findings.

3.3.2. Case 2: Syngas Conversion on ZnCrO Catalysts. Zinc–chromium oxide ($ZnCrO$) catalyst is the first generation industry catalysts for syngas-to-methanol. Experiments show that the catalytic activity is significantly affected by Zn:Cr ratios:^{100–104} the best activity and selectivity is achieved at Zn:Cr = $\sim 1:1$, while the Zn:Cr = 1:2 catalysts yield rather poor activity.¹⁰⁴ The atomic structure of $ZnCrO$ catalysts with the Zn:Cr = 1:1 remains however uncertain.^{104–106} To resolve where and how syngas conversion occurs on $ZnCrO$ catalyst, SSW-NN simulations were performed to obtain the Zn_xCr_yO global PES based on the $ZnCrO$ G-NN potential, which was obtained by exploring different compositions of Zn–Cr–O three element systems.¹⁰⁷ The final G-NN has three-hidden layers (324-80-60-60-1 net), equivalent to 103,743 network parameters in total. The final RMSE of energy and force are 4.3 meV/atom and 0.128 eV/Å respectively.

By using the GMs of different $ZnCrO$ compositions, the thermodynamics phase diagram of Zn–Cr–O is thus constructed, as shown in Figure 9a. It reveals a small stable

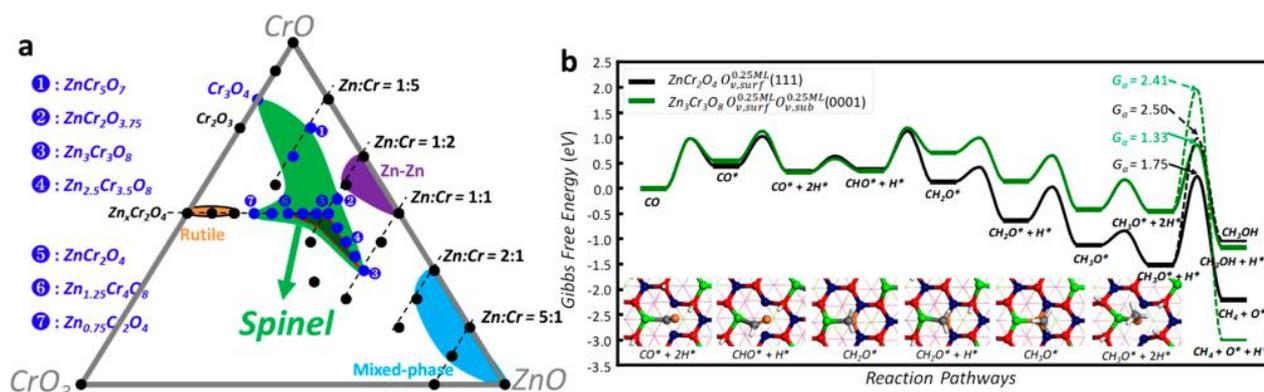


Figure 9. (a) Ternary Zn–Cr–O phase diagram. The green region maps out that the composition with the spinel-type skeleton structure as the global minimum, where blue circles labeled by numbers indicate the composition. (b) Gibbs free-energy reaction profiles for syngas conversion on two ZnCrO catalysts at 573 K and 2.5 MPa syngas ($H_2:CO = 1.5$). The reaction snapshots are shown in the inset of (b), where the color scheme for atoms is as follows: Zn: green; Cr: purple; O: red; O in CO: orange; C: gray and H: white. Figure adapted from ref 107. Copyright 2019 Nature Publishing Group.

composition island (i.e., Zn:Cr:O = 6:6:16~3:8:16), where the oxide alloy crystallizes into a spinel phase. At Zn:Cr = 1:1, a $Zn_3Cr_3O_8$ metastable crystal phase is present, also with the spinel crystal structure, but contains the highest concentration of unusual $[ZnO_6]$ octahedra (O_h) in bulk compared to the other spinel crystals. This subtle structural difference turns out to be critical to affect the syngas conversion activity and selectivity profoundly.

The ZnCrO G-NN potential was utilized for bulk and surface structure determination under reaction conditions and, due to the lack of C and H elements, it cannot be applied to describe the syngas reaction. The efforts to build five-element G-NN potential for catalytic reactions are considerably huge and thus DFT calculations were applied to investigate the syngas conversion reaction on the most stable $Zn_3Cr_3O_8$ and $ZnCr_2O_4$ surfaces determined from G-NN. It was found that the $Zn_3Cr_3O_8$ surface can generate the four-coordinated planar Cr^{2+} cation site on surface that is critical for methanol activity and selectivity. As shown in Figure 9b, the reaction follows $CO \rightarrow CHO \rightarrow CH_2O \rightarrow CH_3O \rightarrow CH_3OH/CH_4$ mechanism, where the rate-determining step is the hydrogenation of CH_3O step, which also determines the selectivity of CH_3OH or CH_4 . The microkinetics simulation further confirms the sharp difference in activity and selectivity observed in experiment.^{104,108}

3.3.3. Case 3: Origin of Amorphous TiO_xH_y for HER. Black titania (TiO_2) exhibit a core–shell structure with the amorphous shell coated on the anatase crystals has been found to exhibit the superior HER activity.^{109–114} The amorphous shell structure which is responsible for the enhanced HER activity is unclear. To resolve the HER active sites on the amorphous TiO_xH_y shell, a Ti–O–H three elements NN potential was constructed to describe the PES of TiO_xH_y system based on large TiO_xH_y global data set with 143 786 structures.¹¹⁵ The network involves two-hidden layers, each with 50 neurons, equivalent to 38 103 network parameters in total. The final RMSE of energy and force are 9.8 meV/atom and 0.22 eV/Å respectively. This TiOH NN potential can replace theoretically the DFT calculations not only in structural identification of TiO_xH_y catalyst but also in catalysis reaction involving water, O_2 , H_2 , and H_2O_2 .

By performing SSW-NN simulation, the thermodynamics phase diagram of TiO_2 bulk and surface in contact with H_2 at

different temperature and pressures can thus be determined quantitatively. Among common anatase surfaces, only the ridged anatase (112) surface was found to reconstruct significantly by surface H atom and a local high H coverage, 0.69 ML, can gradually built up during the surface amorphization with the exposure of 25% Ti_{4c} , 50% Ti_{5c} and 25% Ti_{6c} atoms on surface (Figure 10a). The Ti–O bond length has a wide distribution, from 1.8 to 2.2 Å, as compared with 1.9~2.1 Å on perfect TiO_2 surface. This high H coverage not only renders the black color of the amorphous TiO_2 but also provides a low energy reaction channel for HER: a

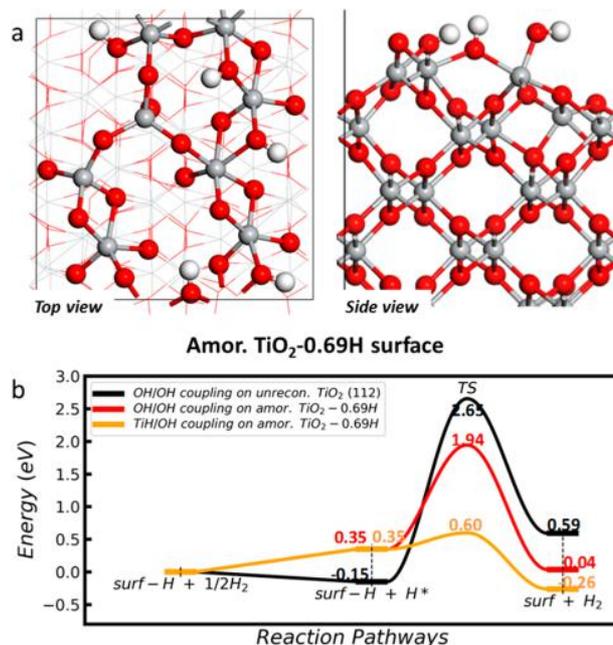


Figure 10. (a) The structure of amorphous TiO_2 surface with H coverage of 0.69 ML from SSW-NN global search. Ti: gray balls; O: red balls; H: white balls (H in reaction: green ball). (b) The energetic profiles of H coupling via the OH/OH coupling and the TiH/OH coupling mechanisms on the pristine (112) surface and amorphous TiO_2 -0.69H. Figure adapted from ref 115. Copyright 2018 American Chemical Society.

transient Ti–H hydride becomes likely to form on the exposed Ti atoms of amorphous surface. The nascent TiH hydride can react readily with the neighboring OH to produce H₂, where the barrier is more than 1 eV lower than the traditional H coupling channel via two surface OH groups (barrier >1.9 eV), as shown in Figure 10b.^{115,116} The TiH/OH coupling pathway for hydrogen production is ~24 orders of magnitude faster than the direct OH/OH coupling, indicating that the TiH/OH mechanism is the only likely reaction for the HER on amorphous TiO₂ material. The experiment also revealed the presence of TiH hydride on amorphous TiO₂, where the characteristic H chemical shift δ value in ¹H nuclear magnetic resonance spectroscopy has a negative peak (−0.6 ppm) and this peak of δ grows with the increase of hydrogenation time.¹¹⁷

4. CONCLUDING REMARKS

The ever-increasing ease to access big data has reshaped our society and also the way of chemical research. In the past decade, a variety of ML techniques have been developed, either based on experimental data or using first-principles results, which demonstrate their great power to solve outstanding challenges in heterogeneous catalysis that are otherwise difficult to treat with traditional methods. While here we retrospect briefly the history of theoretical studies on surface reactions and introduce the main branches in ML methods and applications for catalysis, our main focus is put on the ML-based atomic simulation, which has the great potential to replace DFT for large-scale atomic simulation. As a representative in this field, SSW-NN method implemented in LASP software is elaborated, highlighting the key features of algorithm and the main functionalities. Three cases studies of SSW-NN (i.e., YZrO, ZnCrO, and TiOH systems) for material and reaction simulation are presented to demonstrate the ability of ML for large-scale simulation and the finding of new structures and new reactions.

Apart from the successful applications, the current ML methods do exhibit many obvious shortcomings that need to improve in the future. For example, the limited availability and possible inconsistency are generally present for experimental data, which is a key obstacle for ML to guide experiments; the intrinsic approximation in BEP relation casts doubts on ML-based catalyst and activity prediction using simple descriptors (e.g., the adsorption energy). ML-based atomic simulation using G-NN potentials, although has reached to the stage for wide applications, still suffers from a number of scientific and technical difficulties, as listed in the following.

- (i) The high-cost to build multielement G-NN potentials for materials and reactions. For example, for CuZnAl catalyst for methanol synthesis, it would require a six-element potential, CuZnAlCHO, to explore the phase space of catalyst structures and reactions. The construction of such a G-NN potential is, however, extremely challenging due to the cost to generate the representative PES data;
- (ii) The lack of electronic structure information in the current ML potentials. The electronic structure information, such as atomic charge and density of states, if available with a low-cost, would facilitate greatly to interpret different electronic spectra in experiment data and to understand better how catalysts evolve during reaction;
- (iii) The questionable accuracy for simulating long-range disordering of complex systems, such as glassy materials. There are general concerns on both the data set generation using small unit cells and the use of local atomic descriptors for fitting the energy of long-range disordered structures and the interaction therein.
- (iv) the difficulty to go beyond DFT accuracy. For the great difficulty in generating global data set, DFT calculation remains to be the practical choice for constructing ML potentials. This effectively restricts the accuracy of ML potentials to the level of DFT.
- (v) Overall, there are still ample rooms to design new ML models and invent new ML methods to better cope with the above issues. Luckily, heterogeneous catalysis is complex enough to drive the development of these new technologies for exciting future.

■ AUTHOR INFORMATION

Corresponding Author

Zhi-Pan Liu – Collaborative Innovation Center of Chemistry for Energy Material, Shanghai Key Laboratory of Molecular Catalysis and Innovative Materials, Key Laboratory of Computational Physical Science, Department of Chemistry, Fudan University, Shanghai 200433, China; orcid.org/0000-0002-2906-5217; Email: zpliu@fudan.edu.cn

Author

Sicong Ma – Collaborative Innovation Center of Chemistry for Energy Material, Shanghai Key Laboratory of Molecular Catalysis and Innovative Materials, Key Laboratory of Computational Physical Science, Department of Chemistry, Fudan University, Shanghai 200433, China; orcid.org/0000-0001-5894-5910

Complete contact information is available at:
<https://pubs.acs.org/10.1021/acscatal.0c03472>

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was supported by the National Key Research and Development Program of China (2018YFA0208600), National Science Foundation of China (91945301, 91745201 and 21533001), China Post-Doctoral Science Foundation (2019M661340).

■ REFERENCES

- (1) Ertl, G.; Knözinger, H.; Weitkamp, J. *Handbook of heterogeneous catalysis*; Wiley-VCH: Weinheim, 1997; p 65.
- (2) Liu, Z. P.; Hu, P.; Alavi, A. Catalytic role of gold in gold-based catalysts: A density functional theory study on the CO oxidation on gold. *J. Am. Chem. Soc.* **2002**, *124* (49), 14770–14779.
- (3) Liu, Z. P.; Hu, P. General rules for predicting where a catalytic reaction should occur on metal surfaces: A density functional theory study of C-H and C-O bond breaking/making on flat, stepped, and kinked metal surfaces. *J. Am. Chem. Soc.* **2003**, *125* (7), 1958–1967.
- (4) Erisman, J. W.; Sutton, M. A.; Galloway, J.; Klimont, Z.; Winiwarter, W. How a century of ammonia synthesis changed the world. *Nat. Geosci.* **2008**, *1* (10), 636–639.
- (5) Hellman, A.; Baerends, E. J.; Biczysko, M.; Bligaard, T.; Christensen, C. H.; Clary, D. C.; Dahl, S.; van Harrevelt, R.; Honkala, K.; Jonsson, H.; Kroes, G. J.; Luppi, M.; Manthe, U.; Norskov, J. K.; Olsen, R. A.; Rossmeisl, J.; Skulason, E.; Tautermann, C. S.; Varandas, A. J. C.; Vincent, J. K. Predicting catalysis: Understanding ammonia

- synthesis from first-principles calculations. *J. Phys. Chem. B* **2006**, *110* (36), 17719–17735.
- (6) Honkala, K.; Hellman, A.; Remediakis, I. N.; Logadottir, A.; Carlsson, A.; Dahl, S.; Christensen, C. H.; Norskov, J. K. Ammonia synthesis from first-principles calculations. *Science* **2005**, *307* (5709), 555–558.
- (7) Qian, J.; An, Q.; Fortunelli, A.; Nielsen, R. J.; Goddard, W. A. Reaction Mechanism and Kinetics for Ammonia Synthesis on the Fe(111) Surface. *J. Am. Chem. Soc.* **2018**, *140* (20), 6288–6297.
- (8) Fuller, J.; Fortunelli, A.; Goddard, W. A.; An, Q. Reaction mechanism and kinetics for ammonia synthesis on the Fe(211) reconstructed surface. *Phys. Chem. Chem. Phys.* **2019**, *21* (21), 11444–11454.
- (9) Ertl, G.; Lee, S. B.; Weiss, M. Adsorption Of Nitrogen on Potassium Promoted Fe(111) And (100) Surfaces. *Surf. Sci.* **1982**, *114* (2–3), 527–545.
- (10) Truhlar, D. G.; Garrett, B. C.; Klippenstein, S. J. Current status of transition-state theory. *J. Phys. Chem.* **1996**, *100* (31), 12771–12800.
- (11) Truhlar, D. G.; Hase, W. L.; Hynes, J. T. Current Status Of Transition-State Theory. *J. Phys. Chem.* **1983**, *87* (15), 2664–2682.
- (12) Zhang, C. J.; Liu, Z. P.; Hu, P. Stepwise addition reactions in ammonia synthesis: A first principles study. *J. Chem. Phys.* **2001**, *115* (2), 609–611.
- (13) Norskov, J. K.; Studt, F.; Abild-Pedersen, F.; Bligaard, T. *Fundamental concepts in heterogeneous catalysis*; John Wiley & Sons, 2014; p 265.
- (14) Kresse, G. Ab-Initio Molecular-Dynamics for Liquid-Metals. *J. Non-Cryst. Solids* **1995**, *193*, 222–229.
- (15) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized gradient approximation made simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865.
- (16) Heyd, J.; Scuseria, G. E.; Ernzerhof, M. Hybrid functionals based on a screened Coulomb potential. *J. Chem. Phys.* **2003**, *118* (18), 8207–8215.
- (17) Alavi, A.; Hu, P. J.; Deutsch, T.; Silvestrelli, P. L.; Hutter, J. CO oxidation on Pt(111): An ab initio density functional theory study. *Phys. Rev. Lett.* **1998**, *80* (16), 3650–3653.
- (18) Henkelman, G.; Jonsson, H. A dimer method for finding saddle points on high dimensional potential surfaces using only first derivatives. *J. Chem. Phys.* **1999**, *111* (15), 7010–7022.
- (19) Olsen, R. A.; Kroes, G. J.; Henkelman, G.; Arnaldsson, A.; Jonsson, H. Comparison of methods for finding saddle points without knowledge of the final states. *J. Chem. Phys.* **2004**, *121* (20), 9776–9792.
- (20) Shang, C.; Liu, Z.-P. Constrained Broyden minimization combined with the dimer method for locating transition state of complex reactions. *J. Chem. Theory Comput.* **2010**, *6* (4), 1136–1144.
- (21) Henkelman, G.; Jonsson, H. Improved tangent estimate in the nudged elastic band method for finding minimum energy paths and saddle points. *J. Chem. Phys.* **2000**, *113* (22), 9978–9985.
- (22) Henkelman, G.; Uberuaga, B. P.; Jonsson, H. A climbing image nudged elastic band method for finding saddle points and minimum energy paths. *J. Chem. Phys.* **2000**, *113* (22), 9901–9904.
- (23) Zhang, X.-J.; Shang, C.; Liu, Z.-P. Double-ended surface walking method for pathway building and transition state location of complex reactions. *J. Chem. Theory Comput.* **2013**, *9* (12), 5745–5753.
- (24) Anglada, J. M.; Besalu, E.; Boffil, J. M.; Crehuet, R. Prediction of approximate transition states by Bell-Evans-Polanyi principle: II. Gas phase unimolecular decomposition of methylidioxirane. *J. Comput. Chem.* **1999**, *20* (11), 1130–1137.
- (25) Norskov, J. K.; Bligaard, T.; Logadottir, A.; Bahn, S.; Hansen, L. B.; Bollinger, M.; Bengard, H.; Hammer, B.; Sljivancanin, Z.; Mavrikakis, M.; Xu, Y.; Dahl, S.; Jacobsen, C. J. H. Universality in heterogeneous catalysis. *J. Catal.* **2002**, *209* (2), 275–278.
- (26) Michaelides, A.; Liu, Z. P.; Zhang, C. J.; Alavi, A.; King, D. A.; Hu, P. Identification of general linear relationships between activation energies and enthalpy changes for dissociation reactions at surfaces. *J. Am. Chem. Soc.* **2003**, *125* (13), 3704–3705.
- (27) Pallassana, V.; Neurock, M. Electronic factors governing ethylene hydrogenation and dehydrogenation activity of pseudomorphic Pd-ML/Re(0001), Pd-ML/Ru(0001), Pd(111), and Pd-ML/Au(111) surfaces. *J. Catal.* **2000**, *191* (2), 301–317.
- (28) Hammer, B.; Norskov, J. K. Theoretical surface science and catalysis - Calculations and concepts. *Adv. Catal.* **2000**, *45*, 71–129.
- (29) Calle-Vallejo, F.; Martinez, J. I.; Garcia-Lastra, J. M.; Sautet, P.; Loffreda, D. Fast Prediction of Adsorption Properties for Platinum Nanocatalysts with Generalized Coordination Numbers. *Angew. Chem., Int. Ed.* **2014**, *53* (32), 8316–8319.
- (30) Calle-Vallejo, F.; Tymoczko, J.; Colic, V.; Vu, Q. H.; Pohl, M. D.; Morgenstern, K.; Loffreda, D.; Sautet, P.; Schuhmann, W.; Bandarenka, A. S. Finding optimal surface sites on heterogeneous catalysts by counting nearest neighbors. *Science* **2015**, *350* (6257), 185–189.
- (31) Greeley, J.; Jaramillo, T. F.; Bonde, J.; Chorkendorff, I. B.; Norskov, J. K. Computational high-throughput screening of electrocatalytic materials for hydrogen evolution. *Nat. Mater.* **2006**, *5* (11), 909–913.
- (32) Norskov, J. K.; Bligaard, T.; Rossmeisl, J.; Christensen, C. H. Towards the computational design of solid catalysts. *Nat. Chem.* **2009**, *1* (1), 37–46.
- (33) Reuter, K.; Frenkel, D.; Scheffler, M. The steady state of heterogeneous catalysis, studied by first-principles statistical mechanics. *Phys. Rev. Lett.* **2004**, *93* (11), 116105.
- (34) Tang, Q. L.; Hong, Q. J.; Liu, Z. P. CO₂ fixation into methanol at Cu/ZrO₂ interface from first principles kinetic Monte Carlo. *J. Catal.* **2009**, *263* (1), 114–122.
- (35) Chen, C.; Zuo, Y. X.; Ye, W. K.; Li, X. G.; Deng, Z.; Ong, S. P. A Critical Review of Machine Learning of Energy Materials. *Adv. Energy Mater.* **2020**, *10* (8), 1903242.
- (36) Toyao, T.; Maeno, Z.; Takakusagi, S.; Kamachi, T.; Takigawa, I.; Shimizu, K. Machine Learning for Catalysis Informatics: Recent Applications and Prospects. *ACS Catal.* **2020**, *10* (3), 2260–2297.
- (37) Goldsmith, B. R.; Esterhuizen, J.; Liu, J. X.; Bartel, C. J.; Sutton, C. Machine learning for heterogeneous catalyst design and discovery. *AIChE J.* **2018**, *64* (7), 2311–2323.
- (38) Kitchin, J. R. Machine learning in catalysis. *Nat. Catal.* **2018**, *1* (4), 230–232.
- (39) Huang, S.-D.; Shang, C.; Zhang, X.-J.; Liu, Z.-P. Material discovery by combining stochastic surface walking global optimization with a neural network. *Chem. Sci.* **2017**, *8* (9), 6327–6337.
- (40) Huang, S. D.; Shang, C.; Kang, P. L.; Zhang, X. J.; Liu, Z. P. LASP: Fast global potential energy surface exploration. *WIREs Comput. Mol. Sci.* **2019**, e1415.
- (41) Michie, D.; Spiegelhalter, D. J.; Taylor, C. C. *Machine learning, Neural and Statistical Classification*; Ellis Horwood, University College, London, 1994; p 12.
- (42) Sasaki, M.; Hamada, H.; Kintaichi, Y.; Ito, T. Application Of a Neural-Network To the Analysis Of Catalytic Reactions Analysis Of No Decomposition over Cu/Zsm-5 Zeolite. *Appl. Catal., A* **1995**, *132* (2), 261–270.
- (43) Mohammed, M. L.; Patel, D.; Mbeleck, R.; Niyogi, D.; Sherrington, D. C.; Saha, B. Optimisation of alkene epoxidation catalysed by polymer supported Mo (VI) complexes and application of artificial neural network for the prediction of catalytic performances. *Appl. Catal., A* **2013**, *466*, 142–152.
- (44) Kite, S.; Hattori, T.; Murakami, Y. Estimation Of Catalytic Performance by Neural-Network-Product Distribution In Oxidative Dehydrogenation Of Ethylbenzene. *Appl. Catal., A* **1994**, *114* (2), L173–L178.
- (45) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A. Commentary: The Materials Project: A materials genome approach to accelerating materials innovation. *APL Mater.* **2013**, *1* (1), No. 011002.
- (46) Rahman, M. B. A.; Chaibakhsh, N.; Basri, M.; Salleh, A. B.; Rahman, R. N. Z. R. A Application of Artificial Neural Network for

Yield Prediction of Lipase-Catalyzed Synthesis of Dioctyl Adipate. *Appl. Biochem. Biotechnol.* **2009**, *158* (3), 722–735.

(47) Kim, E.; Huang, K.; Saunders, A.; McCallum, A.; Ceder, G.; Olivetti, E. Materials Synthesis Insights from Scientific Literature via Text Extraction and Machine Learning. *Chem. Mater.* **2017**, *29* (21), 9436–9444.

(48) Baysal, M.; Gunay, M. E.; Yildirim, R. Decision tree analysis of past publications on catalytic steam reforming to develop heuristics for high performance: A statistical review. *Int. J. Hydrogen Energy* **2017**, *42* (1), 243–254.

(49) Gunay, M. E.; Yildirim, R. Knowledge Extraction from Catalysis of the Past: A Case of Selective CO Oxidation over Noble Metal Catalysts between 2000 and 2012. *ChemCatChem* **2013**, *5* (6), 1395–1406.

(50) Gunay, M. E.; Yildirim, R. Neural network Analysis of Selective CO Oxidation over Copper-Based Catalysts for Knowledge Extraction from Published Data in the Literature. *Ind. Eng. Chem. Res.* **2011**, *50* (22), 12488–12500.

(51) Omata, K. Screening of New Additives of Active-Carbon-Supported Heteropoly Acid Catalyst for Friedel-Crafts Reaction by Gaussian Process Regression. *Ind. Eng. Chem. Res.* **2011**, *50* (19), 10948–10954.

(52) Burger, B.; Maffettone, P. M.; Gusev, V. V.; Aitchison, C. M.; Bai, Y.; Wang, X. Y.; Li, X. B.; Alston, B.; Li, B. Y.; Clowes, R.; Rankin, N.; Harris, B.; Sprick, R. S.; Cooper, A. I. A mobile robotic chemist. *Nature* **2020**, *583* (7815), 237–241.

(53) 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.; Norskov, 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.

(54) Liu, X. Y.; Xiao, J. P.; Peng, H. J.; Hong, X.; Chan, K.; Norskov, J. K. Understanding trends in electrochemical carbon dioxide reduction rates. *Nat. Commun.* **2017**, *8*, 15438.

(55) Jinnouchi, R.; Asahi, R. Predicting Catalytic Activity of Nanoparticles by a DFT-Aided Machine-Learning Algorithm. *J. Phys. Chem. Lett.* **2017**, *8* (17), 4279–4283.

(56) Jinnouchi, R.; Hirata, H.; Asahi, R. Extrapolating Energetics on Clusters and Single-Crystal Surfaces to Nanoparticles by Machine-Learning Scheme. *J. Phys. Chem. C* **2017**, *121* (47), 26397–26405.

(57) Tran, K.; Ulissi, Z. W. Active learning across intermetallics to guide discovery of electrocatalysts for CO₂ reduction and H₂ evolution. *Nat. Catal.* **2018**, *1* (9), 696–703.

(58) Liu, X.; Xiao, J.; Peng, H.; Hong, X.; Chan, K.; Norskov, J. K. Understanding trends in electrochemical carbon dioxide reduction rates. *Nat. Commun.* **2017**, *8* (1), 15438.

(59) Zhong, M.; Tran, K.; Min, Y.; Wang, C.; Wang, Z.; Dinh, C.-T.; De Luna, P.; Yu, Z.; Rasouli, A. S.; Brodersen, P.; Sun, S.; Voznyy, O.; Tan, C.-S.; Askerka, M.; Che, F.; Liu, M.; Seifitokaldani, A.; Pang, Y.; Lo, S.-C.; Ip, A.; Ulissi, Z.; Sargent, E. H. Accelerated discovery of CO₂ electrocatalysts using active machine learning. *Nature* **2020**, *581* (7807), 178–183.

(60) Ulissi, Z. W.; Medford, A. J.; Bligaard, T.; Norskov, J. K. To address surface reaction network complexity using scaling relations machine learning and DFT calculations. *Nat. Commun.* **2017**, *8*, 14621.

(61) Toyao, T.; Suzuki, K.; Kikuchi, S.; Takakusagi, S.; Shimizu, K.-i.; Takigawa, I. Toward Effective Utilization of Methane: Machine Learning Prediction of Adsorption Energies on Metal Alloys. *J. Phys. Chem. C* **2018**, *122* (15), 8315–8326.

(62) Wexler, R. B.; Martínez, J. M. P.; Rappe, A. M. Chemical Pressure-Driven Enhancement of the Hydrogen Evolving Activity of Ni₂P from Nonmetal Surface Doping Interpreted via Machine Learning. *J. Am. Chem. Soc.* **2018**, *140* (13), 4678–4683.

(63) Li, Z.; Ma, X.; Xin, H. Feature engineering of machine-learning chemisorption models for catalyst design. *Catal. Today* **2017**, *280*, 232–238.

(64) Li, Z.; Wang, S.; Chin, W. S.; Achenie, L. E.; Xin, H. High-throughput screening of bimetallic catalysts enabled by machine learning. *J. Mater. Chem. A* **2017**, *5* (46), 24131–24138.

(65) Takigawa, I.; Shimizu, K. I.; Tsuda, K.; Takakusagi, S. Machine-learning prediction of the d-band center for metals and bimetals. *RSC Adv.* **2016**, *6* (58), 52587–52595.

(66) Takigawa, I.; Shimizu, K.; Tsuda, K.; Takakusagi, S. Machine learning predictions of factors affecting the activity of heterogeneous metal catalysts. *Nanoinformatics*; Springer: Singapore, 2018; pp 45–64.

(67) Sun, Y. M.; Liao, H. B.; Wang, J. R.; Chen, B.; Sun, S. N.; Ong, S. J. H.; Xi, S. B.; Diao, C. Z.; Du, Y. H.; Wang, J. O.; Breese, M. B. H.; Li, S. Z.; Zhang, H.; Xu, Z. C. J. Covalency competition dominates the water oxidation structure-activity relationship on spinel oxides. *Nat. Catal.* **2020**, *3*, 554–563.

(68) Bartók, A. P.; Gillan, M. J.; Manby, F. R.; Csányi, G. Machine-learning approach for one- and two-body corrections to density functional theory: Applications to molecular and condensed water. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2013**, *88* (5), No. 054104.

(69) Behler, J.; Parrinello, M. Generalized neural-network representation of high-dimensional potential-energy surfaces. *Phys. Rev. Lett.* **2007**, *98* (14), 146401.

(70) Behler, J. Atom-centered symmetry functions for constructing high-dimensional neural network potentials. *J. Chem. Phys.* **2011**, *134* (7), No. 074106.

(71) Behler, J. Representing potential energy surfaces by high-dimensional neural network potentials. *J. Phys.: Condens. Matter* **2014**, *26* (18), 183001–1830024.

(72) Behler, J. First Principles Neural Network Potentials for Reactive Simulations of Large Molecular and Condensed Systems. *Angew. Chem., Int. Ed.* **2017**, *56* (42), 12828–12840.

(73) Bartók, A. P.; Kondor, R.; Csányi, G. On representing chemical environments. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2013**, *87* (18), 184115.

(74) Deringer, V. L.; Csányi, G. Machine learning based interatomic potential for amorphous carbon. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2017**, *95* (9), No. 094203.

(75) Rowe, P.; Csányi, G.; Alfé, D.; Michaelides, A. Development of a machine learning potential for graphene. *Phys. Rev. B: Condens. Matter Mater. Phys.* **2018**, *97* (5), No. 054303.

(76) Boes, J. R.; Kitchin, J. R. Neural network predictions of oxygen interactions on a dynamic Pd surface. *Mol. Simul.* **2017**, *43* (5–6), 346–354.

(77) Rose, M. K.; Borg, A.; Dunphy, J. C.; Mitsui, T.; Ogletree, D. F.; Salmeron, A. Chemisorption of atomic oxygen on Pd(111) studied by STM. *Surf. Sci.* **2004**, *561* (1), 69–78.

(78) Artrith, N.; Hiller, B.; Behler, J. Neural network potentials for metals and oxides—First applications to copper clusters at zinc oxide. *Phys. Status Solidi B* **2013**, *250* (6), 1191–1203.

(79) Shang, C.; Liu, Z.-P. Stochastic surface walking method for structure prediction and pathway searching. *J. Chem. Theory Comput.* **2013**, *9* (3), 1838–1845.

(80) Zhang, X.-J.; Shang, C.; Liu, Z.-P. From atoms to fullerene: stochastic surface walking solution for automated structure prediction of complex material. *J. Chem. Theory Comput.* **2013**, *9* (7), 3252–3260.

(81) Huang, S.-D.; Shang, C.; Kang, P.-L.; Liu, Z.-P. Atomic structure of boron resolved using machine learning and global sampling. *Chem. Sci.* **2018**, *9* (46), 8644–8655.

(82) Li, X.-T.; Chen, L.; Wei, G.-F.; Shang, C.; Liu, Z.-P. Catalytic Selectivity Sharp Increase in Acetylene Semi-hydrogenation on Pd Achieved by Machine Learning Simulation Guided Experiment. *ACS Catal.* **2020**, *10* (17), 9694–9705.

(83) Kirkpatrick, S.; Gelatt, C. D.; Vecchi, M. P. Optimization by simulated annealing. *Science* **1983**, *220* (4598), 671–680.

(84) Gastegger, M.; Marquetand, P. High-Dimensional Neural Network Potentials for Organic Reactions and an Improved Training Algorithm. *J. Chem. Theory Comput.* **2015**, *11* (5), 2187–2198.

- (85) Herr, J. E.; Yao, K.; McIntyre, R.; Toth, D. W.; Parkhill, J. Metadynamics for training neural network model chemistries: A competitive assessment. *J. Chem. Phys.* **2018**, *148* (24), 241710.
- (86) Wales, D. J.; Doye, J. P. Global optimization by basin-hopping and the lowest energy structures of Lennard-Jones clusters containing up to 110 atoms. *J. Phys. Chem. A* **1997**, *101* (28), 5111–5116.
- (87) Jacobsen, T. L.; Jorgensen, M. S.; Hammer, B. On-the-Fly Machine Learning of Atomic Potential in Density Functional Theory Structure Optimization. *Phys. Rev. Lett.* **2018**, *120* (2), No. 026102.
- (88) Artrith, N.; Urban, A.; Ceder, G. Constructing first-principles phase diagrams of amorphous Li_x Si using machine-learning-assisted sampling with an evolutionary algorithm. *J. Chem. Phys.* **2018**, *148* (24), 241711.
- (89) Sivanandam, S.; Deepa, S. In *Introduction to Genetic Algorithms*; Springer, 2008; pp165–209.
- (90) Tong, Q.; Xue, L.; Lv, J.; Wang, Y.; Ma, Y. Accelerating CALYPSO structure prediction by data-driven learning of a potential energy surface. *Faraday Discuss.* **2018**, *211* (0), 31–43.
- (91) Shang, C.; Zhang, X.-J.; Liu, Z.-P. Stochastic surface walking method for crystal structure and phase transition pathway prediction. *Phys. Chem. Chem. Phys.* **2014**, *16* (33), 17845–17856.
- (92) Steele, B. C.; Heinzel, A. *Materials for Fuel-Cell Technologies*. *Nature* **2001**, *414*, 345–352.
- (93) Ryskhewitch, E. *Oxide ceramics: physical chemistry and technology*. *Oxide ceramics: physical chemistry and technology*; Academic Press: New York (NY), 1960; p 350.
- (94) Zakaria, Z.; Abu Hassan, S. H.; Shaari, N.; Yahaya, A. Z.; Boon Kar, Y. A review on recent status and challenges of yttria stabilized zirconia modification to lowering the temperature of solid oxide fuel cells operation. *Int. J. Energy Res.* **2020**, *44* (2), 631–650.
- (95) Liu, T.; Zhang, X. F.; Wang, X. N.; Yu, J. K.; Li, L. A review of zirconia-based solid electrolytes. *Ionics* **2016**, *22* (12), 2249–2262.
- (96) Etsell, T. H.; Flengas, S. N. Electrical properties of solid oxide electrolytes. *Chem. Rev.* **1970**, *70*, 339–376.
- (97) Kondoh, J.; Kawashima, T.; Kikuchi, S.; Tomii, Y.; Ito, Y. Effect of Aging on Yttria-Stabilized Zirconia I. A Study of Its Electrochemical Properties. *J. Electrochem. Soc.* **1998**, *145* (5), 1527–1536.
- (98) Guan, S. H.; Zhang, K. X.; Shang, C.; Liu, Z. P. Stability and anion diffusion kinetics of Yttria-stabilized zirconia resolved from machine learning global potential energy surface exploration. *J. Chem. Phys.* **2020**, *152* (9), No. 094703.
- (99) Guan, S.-H.; Shang, C.; Liu, Z.-P. Resolving the Temperature and Composition Dependence of Ion Conductivity for Yttria-Stabilized Zirconia from Machine Learning Simulation. *J. Phys. Chem. C* **2020**, *124* (28), 15085–15093.
- (100) Waugh, K. Methanol synthesis. *Catal. Today* **1992**, *15* (1), 51–75.
- (101) Molstad, M. C.; Dodge, B. F. Zinc Oxide–Chromium Oxide Catalysts for Methanol Synthesis. *Ind. Eng. Chem.* **1935**, *27* (2), 134–140.
- (102) Errani, E.; Trifiro, F.; Vaccari, A.; Richter, M.; Del Piero, G. Structure and reactivity of Zn-Cr mixed oxides. Role of non-stoichiometry in the catalytic synthesis of methanol. *Catal. Lett.* **1989**, *3* (1), 65–72.
- (103) Bradford, M. C.; Konduru, M. V.; Fuentes, D. X. Preparation, characterization and application of Cr₂O₃/ZnO catalysts for methanol synthesis. *Fuel Process. Technol.* **2003**, *83* (1–3), 11–25.
- (104) Song, H.; Laudenschleger, D.; Carey, J. J.; Ruland, H.; Nolan, M.; Muhler, M. Spinel-Structured ZnCr₂O₄ with Excess Zn Is the Active ZnO/Cr₂O₃ Catalyst for High-Temperature Methanol Synthesis. *ACS Catal.* **2017**, *7* (11), 7610–7622.
- (105) Del Piero, G.; Trifiro, F.; Vaccari, A. Non-stoichiometric Zn–Cr spinel as active phase in the catalytic synthesis of methanol. *J. Chem. Soc., Chem. Commun.* **1984**, No. 10, 656–658.
- (106) Bertoldi, M.; Fubini, B.; Giamello, E.; Busca, G.; Trifiro, F.; Vaccari, A. Structure and reactivity of zinc–chromium mixed oxides. Part 1.—The role of non-stoichiometry on bulk and surface properties. *J. Chem. Soc., Faraday Trans. 1* **1988**, *84* (5), 1405–1421.
- (107) Ma, S.; Huang, S.-D.; Liu, Z.-P. Dynamic coordination of cations and catalytic selectivity on zinc–chromium oxide alloys during syngas conversion. *Nat. Catal.* **2019**, *2* (8), 671–677.
- (108) Cheng, K.; Gu, B.; Liu, X.; Kang, J.; Zhang, Q.; Wang, Y. Direct and highly selective conversion of synthesis gas into lower olefins: design of a bifunctional catalyst combining methanol synthesis and carbon–carbon coupling. *Angew. Chem.* **2016**, *128* (15), 4803–4806.
- (109) Chen, X.; Liu, L.; Huang, F. Black titanium dioxide (TiO₂) nanomaterials. *Chem. Soc. Rev.* **2015**, *44* (7), 1861–1885.
- (110) Chen, X.; Liu, L.; Yu, P. Y.; Mao, S. S. Increasing solar absorption for photocatalysis with black hydrogenated titanium dioxide nanocrystals. *Science* **2011**, *331* (6018), 746–750.
- (111) Naldoni, A.; Allieta, M.; Santangelo, S.; Marelli, M.; Fabbri, F.; Cappelli, S.; Bianchi, C. L.; Psaro, R.; Dal Santo, V. Effect of nature and location of defects on bandgap narrowing in black TiO₂ nanoparticles. *J. Am. Chem. Soc.* **2012**, *134* (18), 7600–7603.
- (112) Yu, X.; Kim, B.; Kim, Y. K. Highly enhanced photoactivity of anatase TiO₂ nanocrystals by controlled hydrogenation-induced surface defects. *ACS Catal.* **2013**, *3* (11), 2479–2486.
- (113) Liu, N.; Schneider, C.; Freitag, D.; Hartmann, M.; Venkatesan, U.; Müller, J.; Spiecker, E.; Schmuki, P. Black TiO₂ nanotubes: cocatalyst-free open-circuit hydrogen generation. *Nano Lett.* **2014**, *14* (6), 3309–3313.
- (114) Lu, H.; Zhao, B.; Pan, R.; Yao, J.; Qiu, J.; Luo, L.; Liu, Y. Safe and facile hydrogenation of commercial Degussa P25 at room temperature with enhanced photocatalytic activity. *RSC Adv.* **2014**, *4* (3), 1128–1132.
- (115) Ma, S.; Huang, S.-D.; Fang, Y.-H.; Liu, Z.-P. TiH Hydride Formed on Amorphous Black Titania: Unprecedented Active Species for Photocatalytic Hydrogen Evolution. *ACS Catal.* **2018**, *8* (10), 9711–9721.
- (116) Aschauer, U.; Selloni, A. Hydrogen interaction with the anatase TiO₂ (101) surface. *Phys. Chem. Chem. Phys.* **2012**, *14* (48), 16595–16602.
- (117) Guo, Y.; Chen, S.; Yu, Y.; Tian, H.; Zhao, Y.-L.; Ren, J.-C.; Huang, C.; Bian, H.; Huang, M.; An, L.; et al. Hydrogen-Location-Sensitive Modulation of the Redox Reactivity for Oxygen-deficient TiO₂. *J. Am. Chem. Soc.* **2019**, *141* (21), 8407–8411.