Accelerating Catalyst Discovery via Ab Initio Machine Learning

Zheng Li

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
in
Chemical Engineering

Hongliang Xin, Chair
Luke E. Achenie
Sanket A. Deshmukh
Amanda J. Morris

October 23, 2019
Blacksburg, Virginia

Keywords: Computational catalysis, Density functional theory, Machine learning

Copyright 2019, Zheng Li
In recent decades, machine learning techniques have received an explosion of interest in the domain of high-throughput materials discovery, which is largely attributed to the fast-growing development of quantum-chemical methods and learning algorithms. Nevertheless, machine learning for catalysis is still at its initial stage due to our insufficient knowledge of the structure-property relationships. In this regard, we demonstrate a holistic machine-learning framework as surrogate models for the expensive density functional theory to facilitate the discovery of high-performance catalysts. The framework, which integrates the descriptor-based kinetic analysis, material fingerprinting and machine learning algorithms, can rapidly explore a broad range of materials space with enormous compositional and configurational degrees of freedom prior to the expensive quantum-chemical calculations and/or experimental testing. Importantly, advanced machine learning approaches (e.g., global sensitivity analysis, principal component analysis, and exploratory analysis) can be utilized to shed light on the underlying physical factors governing the catalytic activity on a diverse type of catalytic materials with different applications. Chapter 1 introduces some basic concepts and knowledge relating to the computational catalyst design. Chapter 2 and Chapter 3 demonstrate the methodology to construct the machine-learning models for bimetallic catalysts. In Chapter 4, the multi-functionality of the machine-learning models is illustrated to understand the metalloporphyrin’s underlying structure-property relationships. In Chapter 5, an uncertainty-guided machine learning strategy is introduced to tackle the challenge of data deficiency for perovskite electrode materials design in the electrochemical water splitting cell.
Acceralating Catalyst Discovery via Ab Initio Machine Learning

Zheng Li

(GENERAL AUDIENCE ABSTRACT)

Machine learning and deep learning techniques have revolutionized a range of industries in recent years and have huge potential to improve every aspect of our daily lives. Essentially, machine learning provides algorithms the ability to automatically discover the hidden patterns of data without being explicitly programmed. Because of this, machine learning models have gained huge successes in applications such as website recommendation systems, online fraud detection, robotic technologies, image recognition, etc. Nevertheless, implementing machine learning techniques in the field of catalyst design remains difficult due to 2 primary challenges. The first challenge is our insufficient knowledge about the structure-property relationships for diverse material systems. Typically, developing a physically intuitive material feature method requests in-depth expert knowledge about the underlying physics of the material system and it is always an active field. The second challenge is the lack of training data in academic research. In many cases, collecting a sufficient amount of training data is not always feasible due to the limitation of computational/experimental resources. Subsequently, the machine learning model optimized with small data tends to be over-fitted and could provide biased predictions with huge uncertainties. To address the above-mentioned challenges, this thesis focuses on the development of robust feature methods and strategies for a variety of catalyst systems using the density functional theory (DFT) calculations. Through the case studies in the chapters, we demonstrate that machine learning techniques have huge potential to facilitate the discovery process of promising energy materials.
First of all, I would like to express my sincere gratitude and appreciation to my advisor Prof. Hongliang Xin, for his continuous guidance and support of my Ph.D research and study. More importantly, thanks my advisor for leading me to this fascinating field and I believe the knowledge I gained would have huge benefits for my future career. Besides my advisor, I would like to thank the rest of my committee: Prof. Luke E. Achenie, Prof. Sanket A. Deshmukh, and Dr. Amanda J. Morris, for their valuable suggestions and comments. All the suggestions are great driving forces for me to make continuous improvement. My sincere thanks also goes to all my workmates in the computational catalysis group, who provided me technical assistants and suggestions during my research and studies. Also, I would like thank the Department of Chemical Engineering and Advanced Research Computing (ARC) at Virginia Tech, for providing me the educational and computational resources to enable me to complete my study and researches. Last but not the least, I would like to express my deepest gratitude to my parents Chengqing Li and Guifeng Ba, for giving birth to me and supporting me both spiritually and physically throughout my life. Also, I would like to express my deepest appreciation to my wife Qinghe(Angela) Zheng, for loving me and being understanding with me.
## Contents

List of Figures v

1 Fundamental concepts in computational catalysis 1

1.1 Introduction 1

1.2 Density functional theory 2

1.3 DFT in electrocatalysis 4

1.3.1 Gibbs free energy diagram 6

1.3.2 Quantum and thermal corrections 9

1.3.3 Entropy corrections 11

1.4 Electronic factors in catalysis 13

2 Feature engineering of machine-learning models for alloy electrocatalysts 18

2.1 Introduction 18

2.2 Computational methods 20

2.3 Results and discussion 21

2.3.1 General strategy for machine-learning accelerated catalyst design 21

2.3.2 Numerical fingerprints of metal surfaces 23

2.3.3 Neural network configuration 26
<table>
<thead>
<tr>
<th>Section</th>
<th>Title</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.3.4</td>
<td>Machine-learning models for catalyst prediction</td>
<td>28</td>
</tr>
<tr>
<td>2.4</td>
<td>Conclusions</td>
<td>30</td>
</tr>
<tr>
<td>3</td>
<td>Machine-learning models of bimetallic electrocatalysts</td>
<td>32</td>
</tr>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>32</td>
</tr>
<tr>
<td>3.2</td>
<td>Computational methods</td>
<td>34</td>
</tr>
<tr>
<td>3.3</td>
<td>Results and discussion</td>
<td>35</td>
</tr>
<tr>
<td>3.3.1</td>
<td>Kinetic analysis</td>
<td>35</td>
</tr>
<tr>
<td>3.3.2</td>
<td>Machine-learning model development</td>
<td>38</td>
</tr>
<tr>
<td>3.3.3</td>
<td>Model prediction of bimetallic catalysts</td>
<td>43</td>
</tr>
<tr>
<td>3.3.4</td>
<td>Validation of machine-learning models</td>
<td>45</td>
</tr>
<tr>
<td>3.3.5</td>
<td>Feature importance analysis</td>
<td>45</td>
</tr>
<tr>
<td>3.4</td>
<td>Conclusions</td>
<td>47</td>
</tr>
<tr>
<td>4</td>
<td>Machine learning energy gaps of metalloporphyrins</td>
<td>51</td>
</tr>
<tr>
<td>4.1</td>
<td>Introduction</td>
<td>51</td>
</tr>
<tr>
<td>4.2</td>
<td>Computational methods</td>
<td>54</td>
</tr>
<tr>
<td>4.2.1</td>
<td>DFT specification</td>
<td>54</td>
</tr>
<tr>
<td>4.2.2</td>
<td>Dataset description</td>
<td>54</td>
</tr>
<tr>
<td>4.3</td>
<td>Results and discussion</td>
<td>56</td>
</tr>
<tr>
<td>4.3.1</td>
<td>Machine-learning model development</td>
<td>56</td>
</tr>
</tbody>
</table>
4.3.2 Model prediction of energy gaps .............................. 60
4.3.3 Principal component analysis ............................... 63
4.3.4 Model interpretation ........................................... 63
4.4 Conclusions ......................................................... 66

5 Adaptive machine learning for perovskite electrocatalyst design 67
5.1 Introduction ......................................................... 67
5.2 Computational methods .......................................... 70
5.3 Results and discussion ........................................... 71
  5.3.1 Adaptive learning strategy ................................. 71
  5.3.2 Kinetic analysis ............................................... 72
  5.3.3 Feature engineering ........................................... 73
  5.3.4 Model training .................................................. 75
  5.3.5 Adaptive learning ............................................. 76
  5.3.6 Model screening of perovskite electrocatalysts ........... 79
  5.3.7 Molecular orbital insights ................................. 82
5.4 Conclusions ......................................................... 84

References 85
List of Figures

1.1 Schematic of DFT self-consistency loop ........................................ 5
1.2 Schematic display of the interaction between molecular orbitals of two $H$ atoms 15
1.3 Electronic band structure diagrams of solid material .......................... 16

2.1 Investigated configurations of {100}-terminated alloy models ............... 21
2.2 Schematic representation of the machine-learning accelerated catalyst design approach ........................................................... 22
2.3 Feature engineering in machine-learning chemisorption models for metal catalysis ................................................................. 24
2.4 Neural network model architecture determination ............................ 27
2.5 Evaluation of machine learning model prediction with DFT calculations . 28
2.6 Limiting potential diagram of $CO_2$ electroreduction .......................... 30

3.1 Schematic display of the advantage of alloy-based catalyst’ bi-functionality in catalysis process .................................................... 33
3.2 Theoretical limiting potential (V vs. RHE) of methanol electro-oxidation as a function of $CO*$ and $OH*$ free formation energies relative to $Pt(111)$ .... 37
3.3 {111}-terminated bimetallic surfaces for the machine-learning model development ................................................................. 39
3.4 Exploratory analysis of training data .................................................. 40
3.5 Neural network training performance evaluation ........................................... 42
3.6 Comparison to other regression algorithms in training and test of *CO adsorption energies on model alloys .......................................................... 44
3.7 High-throughput screening of novel bimetallic alloys enabled by neural network model .......................................................... 48
3.8 Comparison between model-predicted and DFT calculated adsorption energies for CO* and OH* on 40 randomly selected A@A₂B₂ type bimetallic catalysts .......................................................... 49
3.9 Feature importance scores for CO* and OH* models based on the sensitivity analysis of neural network models .......................................................... 50
4.1 Conceptual visualization of Porphyrin’s structure-property relationship enabled my machine-learning approach .......................................................... 53
4.2 Structural labeling scheme of porphyrins with varying ligand substitution and metal chelation .......................................................... 55
4.3 Schematics of molecular graph representations .......................................................... 58
4.4 Parity plots for DFT-calculated vs. machine-learning model prediction of energy gaps with molecular graph representations .......................................................... 62
4.5 Two-dimensional principal component analysis (PCA) of different molecular descriptors of porphyrins .......................................................... 64
4.6 Interpretation of machine learning models using global sensitivity analysis .......................................................... 65
5.1 Schematic of the adaptive machine-learning workflow .......................................................... 70
Chapter 1

Fundamental concepts in computational catalysis

1.1 Introduction

Electrocatalysis is considered as an efficient and environmentally friendly solution to the global increasing energy consumption and the industries’ excessive dependence on fossil fuels. As the overall efficiency of electrocatalysis is primarily determined by the rate of a chemical reaction on electrode material surfaces, designing an optimal catalytic electrode material that can modify the surface chemical reaction rate or equilibrium is becoming the central theme of modern catalyst study. However, the search for high performance catalytic materials through the trail-and-error experimentations is extremely challenge due to the structural versatility and vast chemical space. In recent years, the emerge of accurate first-principles density functional theory (DFT) methods largely facilitates the discovery process of novel catalytic materials. More specifically, DFT provides us a computational tool that can be used to approximate a range of materials’ electronic properties. Under the guidance of the first-principle material properties, we can assess the material’s general catalytic activities for different chemical reaction systems through the quantitative analysis such as thermodynamic and kinetic modeling. In this chapter, we briefly introduce some fundamental theory and principles that are involved in the high-throughput materials calculations such as den-
sity functional theory (DFT), Gibbs free energy diagram and metal’s electronic structure descriptors. Understanding the basic concepts and knowledge are extremely important for us to build high-performance machine learning models with respect to different chemical reaction mechanisms.

1.2 Density functional theory

Density functional theory (DFT), a quantum mechanical simulation method, is widely used nowadays to investigate the ground-state electronic structures of materials in the field of physical science. The theoretical framework of DFT involves the first and second Hohenberg-Kohn theorems and the Kohn-Sham method according to the Local Density Approximation (LDA) for the exchange correlation energy. Here, we provide a brief introduction of the core concepts of density functional theory and the implementation methodology in the practical computational simulations.

The Thomas-Fermi model reported in 1928 is considered as the embryonic form of density functional theory. However, this model is erroneous because it completely ignores the electron correlation effect for the exchange energy. In 1964, a new DFT system is reported by Hohenberg and Kohn. Instead of explicitly solve the many-body wavefunction, the Hohenberg and Kohn substitute the complicated wavefunction with a functional of electron density. Therefore, the parameter space is significantly reduced. The new DFT system of Hohenberg and Kohn involves two important theorems that make the foundation of modern DFT theory. In the first theorem, they provide the truth that there exists one-to-one mapping relationships between the ground state energy of material system and the electron density. While in the second theorem, they move further step to prove that the quantity of ground state energy could be obtained by minimizing the energy functional with respect
to the electron density. Even though the H-K theorems had made an extraordinary contribution by identifying a key variable (electron density) of the electron energy functional, it fails to determine the explicit correlations for potential modeling. In another word, the mathematical relationship between energy functional and electron density remains unknown. Only one year later, the Kohn-Sham paradigm[101] is published which allowing DFT to become available in the quantitative calculations in the researches of solid state physics and computational chemistry. Kohn-Sham method successfully transforms the multi-electron problem into a non-interaction system where the electrons are effected by both the external potential and the Coulomb interactions between the electrons such as the exchange and correlation interactions. In this way, the unsolvable wavefunction is represented by a Slater determinant of orbitals which enables the quantification of the kinetic energy functional. It is noteworthy that solving the exchange correlation interactions is the top challenge in the KS-DFT and the explicit rigorous solution for the energy term is remains unknown. To approximate the exchange correlation interactions, the Local-Density Approximation (LDA) according to the uniform electron gas model by Thomas-Fermi theory[179] is applied. In the last a few decades, an enormous research efforts have been devoted to improve the accuracy of LDA-based approximation methods for the exchange-correlation interaction. Now the DFT is becoming a well known electronic structure calculation method which provides a balance between computational cost and accuracy. However DFT still has huge deficient in describing the strongly correlated systems and the dispersion systems such as semiconductor type materials. As a consequence, the DFT development is still an ongoing process.

The system’s ground state energy according to Kohn-Sham scheme can be written as:

$$E(n(r)) = \int v_{\text{ext}}(r)n(r)dr + F[n(r)], \quad (1.1)$$
where \( v_{\text{ext}}(r) \) denotes the external potential from the nuclei and \( n(r) \) indicates the electronic density function at 3-dimensional spatial space. \( F[n(r)] \) is an universal functional which contains the contributions of non-interacting kinetic energy \( T_0[n(r)] \), classical Coulomb interaction and a non-classical self-interaction correction namely exchange-correlation interaction as show below:

\[
F[n(r)] = T_0[n(r)] + \frac{1}{2} \int \int \frac{n(r)n(r')}{|r - r'|} dr' dr + \int n(r)\varepsilon_{xc}(n(r)) dr.
\] (1.2)

Here \( \varepsilon_{xc}(n(r)) \) is the per-electron exchange and correlation energy which can be approximated in numerous ways such as the Local Density Approximation (LDA)\[27\] or more sophisticated methods such Generalized gradient approximations(GGA)\[145\]. The KS-DFT scheme suggests that the ground state energy of a system can be found if the appropriate electron density is identified. In practical calculations, the ground state electronic properties of a system is approximated by an iterative self-consistency loop where the electron density \( n(r) \) is propagated though the process. The self-consistency process is repeated until the change of electron density between the current and previous iterations is smaller than a specific threshold. Figure 1.1 below shows the schematic of DFT self-consistency mechanism.

### 1.3 DFT in electrocatalysis

In this chapter, we briefly introduce the theoretical basis of electrocatalysis and the methodology to apply the DFT-based computational method to investigate the chemical reaction mechanism on the electrode materials. Electrocatalysis is defined as the heterogeneous catalysis on the electrode materials. It facilitates the conversion between the chemical energy and electrical in the electron chemical cells. Designing promising electrode materials with high
1.3. DFT in electrocatalysis

Catalytic activity and selectivity towards desired product is of vital importance in the study of electrocatalysis. However with limited knowledge of the explicit relationship between the materials’ physicochemical properties and the corresponding kinetic mechanism of chemical reactions, the research and development process of high performance electrocatalysts is extremely challenging. In recent years, the electronic structure calculations, in particular the density functional theory (DFT), are widely used to determine the structure-property relationship of a variety of materials by estimating the elementary chemical reaction energetics. A number of successful theory-guided case studies can be found in the applications of electrocatalysts design. For instance, the oxygen reduction reaction (ORR) and hydrogen oxidation reaction (HOR) on the cathode and anode side of a hydrogen-oxygen fuel cell. And the hydrogen evolution reaction (HER) and oxygen evolution reaction (OER) at the cathode and anode side of electrolytic water-splitting cell, respectively.

The central theme of DFT-based computational catalyst design is to minimize the overpotential of a chemical reaction on the catalyst surface by identifying the optimal catalyst with...
specific structure and compositions. Note that overpotential is an electrochemical property that describing the potential difference between the electric potential and the thermodynamically determined potential to drive an electrochemical reaction at a certain rate.\cite{9} In this work, the overpotential can be understood as a respond variable of a multivariate function that are dependent on a variety of properties such as the structure coordinates, composition, reaction pathways, temperature, applied potential, pH, etc. Typically, larger overpotential indicates that extra energy is required to drive the reaction become thermodynamically favorable. However calculating overpotential is not a trivial task, because it requires us to have a clear picture of the free energy diagram for an electrochemical reaction at different applied potentials. Therefore, acquiring the free energy diagram is becoming the first and most essential step for calculating the overpotential.\cite{136}

1.3.1 Gibbs free energy diagram

Most of electrocatalytic reactions involve the transfer of protons and electrons at each elementary step in the reaction mechanism as show in the following model reaction:

\[ XH \rightleftharpoons X^* + H_{aq}^+ + e^-, \]  \hspace{1cm} (1.3)

where \( \ast \) denotes the adsorbed molecular specie on the catalyst surface. To calculate the change of free energy of above elementary step, we need to calculate the Gibbs free energies that are associated with all the components including the reactants and products. Gibbs free energy is a thermodynamic property describing the maximum amount of non-expansion work that a thermodynamically closed system can conduct to the surroundings.\cite{4} In a chemical reaction system, the change in free energy from the reactants to the products is usually a key indicator to estimate whether a reaction is spontaneous \((\Delta G < 0)\) in the forward or reverse
direction, with respect to equilibrium state ($\Delta G = 0$).

The equation below shows the free energy equation of above reaction step:

$$\Delta G = G(X^*) + G(H^+_{aq}) + G(e^-) - G(XH).$$  \hspace{1cm} (1.4)

Here $G(X^*)$ denotes the Gibbs free energy of the adsorbed specie on the catalyst and $G(XH)$ denotes the free energy of the molecular specie. It is noteworthy that calculating the free energy of electron and proton pair for different electrochemical systems is extremely challenge. To tackle the challenge, the Computational Hydrogen Electrode (CHE) model is widely used by computational catalysis society. The CHE model uses the standard hydrogen electrode (SHE) as reference and assumes a standard and equilibrium conditions between the $H_2$ gas and 1 M $H^+$ concentration liquid ($pH = 0$) at 1 bar and 25 °C. Therefore, the chemical potential of the proton and electron pair is assumed to have the same magnitude to the potential of $H_2$ gas at standard conditions (0 V vs. SHE and pH = 0) as show in the following:

$$\mu(H^+_{aq}) + \mu(e^-) = \frac{1}{2}\mu(H_2(g)).$$  \hspace{1cm} (1.5)

To consider the applied potential and pH effect, we need to add another 2 energy terms to the free energy equation, which including $eU$ ($e$ is the electron charge and $U$ is the applied potential vs. SHE) and $k_B T \log(10)pH$ ($k_B$ is the Boltzmann constant and $T$ is the temperature). Typically the pH effect is very small compared to other terms and it is usually ignored during the practical DFT free energy calculations. Even though the CHE model significantly simply our computational work, it has limitations due to the assumptions. The primary limitation of the CHE model is that it doesn’t take into account the electrostatic
field and molecule solvation effects to the adsorbed intermediates in the electrochemical cell. This may causes the biased estimation of the adsorption energies for the large adsorbed species with strong dipoles.

In the following case study, we show the detailed calculation procedures for investigating the free energy diagram of the oxygen evolution reaction (OER) on metallic catalyst surface. The most favorable 4-step OER reaction pathway and the corresponding free energy at each elementary step is show below:

\[
\text{Step\#1 : } H_2O(l) + * \rightleftharpoons OH^* + H_{aq}^+ + e^- \quad (1.6)
\]

\[
\Delta G_1 = \Delta G_{OH^*} \quad (1.7)
\]

\[
\text{Step\#2 : } OH^* \rightleftharpoons O^* + H_{aq}^+ + e^- \quad (1.8)
\]

\[
\Delta G_2 = \Delta G_{O^*} - \Delta G_{OH^*} \quad (1.9)
\]

\[
\text{Step\#3 : } O^* + H_2O(l) \rightleftharpoons OOH^* + H_{aq}^+ + e^- \quad (1.10)
\]

\[
\Delta G_3 = \Delta G_{OOH^*} - \Delta G_{O^*} \quad (1.11)
\]
In this case, the Gibbs free energies of the protons and electrons are 0 due to the assumption of CHE model ($\Delta G_{H_2(g)} = 0$) and the value of $\Delta G_{O_2}$ (4.92 eV) can be obtained from the thermodynamic equilibrium equation of the overall OER equation. Therefore, in order to get the free energy diagram, DFT is only required to calculate the free energies terms that are related with the adsorbates such as $\Delta G_{O^*}$, $\Delta G_{OH^*}$ and $\Delta G_{OOH^*}$. The energy value obtained from the DFT calculate is called ”ground state” energy of a quantum system at $T = 0K$, which is the (local) minimum potential energy where a classical mechanical system is reaching equilibrium state. However a (real) quantum system is generally impossible to stay at rest at such a (local) minimum energy according to the principle of Heisenberg uncertainty.[169] In order to attain a reasonable estimation of the Gibbs free energy of a system, we need to include some additional correction terms (e.g., quantum and thermal corrections and entropy corrections) to the DFT-calculated ground state energy. In the following 2 sections, we will provide some fundamental knowledge of these 2 types of primary corrections.

1.3.2 Quantum and thermal corrections

In this section, we discuss some fundamentals that are related to the quantum and thermal corrections for the practical DFT calculations. The Heisenberg theorem describes a general idea that the uncertainty of a particle’s position $\Delta x$ has direct correlation with the uncertainty of particle’s momentum $\Delta p$ as show below:
\[ \Delta x \ast \Delta p \geq \frac{\hbar}{2}. \quad (1.14) \]

According to Heisenberg theorem, the adsorbed particle on the catalyst surface is never at rest in a specific perfect location. The harmonic vibrations are widely used to approximate the energy fluctuations around the minimum of local potential energy. The formula below describes the quantized energy \( E_n \) of a harmonic oscillator quantum system:

\[ E_n = E_{\text{pot}}^{\text{min}} + (n + \frac{1}{2})\hbar v_i, \quad (1.15) \]

where \( E_{\text{pot}}^{\text{min}} \) is the local minimum of potential energy, \( v_i \) is the vibrational frequency, \( \hbar \) is the Planck’s constant and \( n \) is the quantum number. Note that energy level is the lowest at the ground state (\( n = 0 \)). In practice, we calculate the sum of all the vibrational mode contributions, namely Zero Point Energy correction, by taking account of all the degree of freedom (e.g., bond stretching and bending) for correcting the ground state energy. The Zero Point Energy correction is given by:

\[ ZPE = \sum_{i=1}^{N_{\text{modes}}} ZPE_i = \sum_{i=1}^{N_{\text{modes}}} \frac{1}{2} \hbar v_i. \quad (1.16) \]

According to equation 1.12, the ZPE (eV) is linearly dependent on the vibrational frequency \( v_i \). In another word, the molecule with large vibrational frequency would have significant energy correction contribution to the ground state energy. Therefore, the inclusion of ZPE correction is extremely important for particular molecules such as \( \text{H}_2 \) (4395 cm\(^{-1}\)).

Besides the ZEP-corrected potential energy, the system’s internal energy \( U(T) \) also contains another thermal energy term that is dependent on temperature \( T(K) \). The thermal energy
1.3. DFT in electrocatalysis

describes the amount of heat that needs to be transferred to heat up a system, thus increasing
the overall kinetic and potential energy. It is calculated by integrating the heat capacity
$C_p(T)$ up to the temperature. Nevertheless, the value $C_p(T)$ is very small in regardless of
the temperature, thus allowing us to reduce the computational complexity by ignoring this
contribution since the $C_p(T)$ varies with temperature. Therefore, the overall internal energy
is shown as following:

$$U(T) = E_0 + ZPE + \int_{T=0}^{T} C_p(T')dT'$$

(1.17)

1.3.3 Entropy corrections

We end this chapter by introducing another key thermodynamic quantity, known as entropy $S$, toward the Gibbs free energy. According to the second law of thermodynamics, a spontaneous process will increase the entropy of the universe. Typically, entropy is defined as the number of accessible quantum states and it describes the level of randomness/disorder of a system. More specifically, if the system’s temperature is increased, the kinetic energy is also increased, which allowing extra quantum states becoming accessible at higher potential energy regions of the potential surface.[136] The entropy of an atom or molecule can be quantitatively determined as following:

$$S = k_B ln(W),$$

(1.18)

where $k_B$ is the Boltzmann constant ($1.38 \times 10^{-23} \text{ J/K}$), and $W$ denotes the number accessible states. In the study of heterogeneous catalysis, we are particular interested in the entropy that are associated with the molecules with the gas phase and adsorption phase. A
non-interacting gas phase molecule typically has 4 entropy contributions as show below:

\[
S_{tot} = S_{trans} + S_{rot} + S_{vib} + S_{el},
\]

which are the contributions from translational, rotational, vibrational and electronic, corresponding to different quantum states of a variety of degrees of freedom. When considering a process where a gas phase molecule is adsorbed on a catalyst’s surface, the entropic contributions arise primarily from the translational modes are lost. The translational and rotational degrees of freedom are transformed to the vibrational modes. For simplicity, the entropy contribution with frequency of \(v_i\), to the standard Gibbs free energy can be roughly approximated as below:

\[
S_{v_i}^0 = -k_B \ln(1 - e^{-\frac{h v_i}{k_B T}}).
\]

\[
G_{v_i}^0 = k_B T \ln(1 - e^{-\frac{h v_i}{k_B T}}).
\]

According to equation 1.16, we know the standard entropy depends on the vibrational frequency of the adsorbed molecule. However, in practice, the adsorption entropic contributions to the Gibbs free energy are relatively small compared to the entropic contributions from the gas phase molecules unless the adsorbate has a large number of low frequency vibrational modes. Therefore, we can roughly assume that the gas molecules would lose all the entropy during the adsorption process. Besides the vibrational entropy, the configurational entropy, which depends on the adsorbates’ surface coverage (\(\theta\)), also plays a very important role toward the Gibbs free energy. For a more detailed description of the configurational entropy, see chapter 3 of Nørskov et al. (2012). [136]
1.4 Electronic factors in catalysis

Understanding the chemical bond formed between the reaction intermediates and the catalyst surfaces is one of the primary focuses of computational catalyst design. Typically, the strengths of the chemical bonds are quantified by the adsorption energy values. In many cases, the bonding energies are crucial activity descriptors for a variety of chemical reactions. It is noteworthy that the adsorption process of a reaction intermediate is a complex physical phenomenon, which involves the interactions of the electronic structures between the adsorbed molecules and the solid surface. To build a physically meaningful machine learning model, it is important for us to have a basic understanding of the electronic structure of solid material and its relationship to the adsorption properties during catalysis process. Understanding these basic fundamentals is also very important for us to draw physical insights using the language of molecular orbital theory through the machine learning analysis. In this chapter, we briefly introduce some basic concepts of electronic structure factors such as solid electronic structure, the interaction of molecular orbitals and the electronic band diagrams.

The heterogeneous catalysts are usually solid materials with different types of crystal structures (e.g., metal, metal oxide). Each solid material is composed of a number of atoms, which are arranged with different alignments, namely surface terminations. As the atoms are getting closer in distance within a crystal structure, the wave functions that describing the outermost atomic states of the atoms begin to overlap to form a composite states. The interaction of atomic states would give rise to a split of the energy levels into the bonding and anti-bonding states. According to molecular orbital theory, the Schrödinger equation can be used to approximate the quantity of the interactions between these atomic energy levels with the composite Hamiltonian. Hamiltonian, as an operator for the Schrödinger
equation, involves a kinetic term and a Coulombic term. For illustration, the Hamiltonian for two separated identical atoms is shown as following:

\[ H = \left( -\frac{\hbar^2}{2m_e} \right) \nabla^2 - \frac{e^2}{4\pi\epsilon_0|r - r_0|} \]  \quad (1.22)

where \( \hbar \) denotes the Planck’s constant, \( m_e \) is the electron mass, \( \epsilon_0 \) denotes the permittivity of the vacuum and \( r_0 \) is the center of Coulomb interaction. To calculate the energy levels (\( \varepsilon \)) of a pair of interacted atoms, we need to know the normalized eigenfunctions that are associated with the electronic states of two non-interacting systems, denoted as \( \psi_A \) and \( \psi_B \) respectively. The eigenfunctions of the composition state is assumed to have a linear correlation \( (\psi_{AB} = c_1\psi_A + c_2\psi_B) \) with \( A \) and \( B \) atoms. Besides, the overlapping matrix \( (S_{AB}) \) and the coupling elements \( (V_{AB}) \) are also requested for solving the Schrödinger equation. \( S \) represents the matrix element for the eigenfunctions for the individual atom \( A \) or \( B \). While \( V \) is the hopping matrix element, which indicates the probability that an electron hops to the neighboring atoms in the composite atom system. Note that \( S \) and \( V \) are functions of the atoms’ eigenfunctions \( \psi \). See chapter 12 of Nørskov et al. (2012) [136] for detailed mathematical formulas.

By solving the Schrödinger equation \( H\psi_{AB} = \epsilon\psi_{AB} \), the energy level for the composite state can be determined. Note that the overlap terms between atomic states are very small and this term is ignored. The simplified expression of the energy state is shown as below:

\[ \epsilon = \frac{(\epsilon_A + \epsilon_B) - 2SV \pm \sqrt{(\epsilon_A - \epsilon_B)^2 + 4V^2}}{2} \]  \quad (1.23)

For a simple case of homo-nuclear atoms system \( (\epsilon_A = \epsilon_B = \epsilon) \), above energy expression can be simplified as below:
1.4. Electronic factors in catalysis

\[ \epsilon_\pm = \epsilon \pm V - SV, \]  \hspace{1cm} \text{(1.24)}

which suggests that both of the energy levels of the individual atom are elevated by a repulsive term with the quantity of \(|SV|\). And the subsequent energy levels are separated by a magnitude of 2 times of the coupling matrix element \(2|V|\). According to the Pauli principle, the up-shift of the energy level for both atomic energy levels is attributed to the Pauli repulsion and the separation of the 2 energy levels is attributed from the hybridization effect. The interaction of atomic levels leads to a split of bonding and an anti-bonding state, as shown in Figure 1.2.

![Diagram of molecular orbitals](image)

**Figure 1.2:** Schematic display of the interaction between molecular orbitals of two \(H\) atoms.

Typically, the distance between the anti-bonding and bonding orbital, as shown in Fig. 1.2, is determined by the extend of orbital interaction. While the bonding strength between the two atoms, on the other hand, is determined by the degree of electron occupancy on the molecular orbitals. Fig. 1.2 describes the molecular orbitals of a very simple 2-atoms system.
However, solid materials usually contain a large number of atoms and the corresponding interactions between all the atoms’ orbitals are much complicated. A band diagram as shown in Fig. 1.3, is widely applied to quantitatively describes the degenerated energy levels and the associating orbital-resolved partial density of states for the solid material structure.

![Figure 1.3: Electronic band structure diagrams of solid material. (a) Band diagrams for the cubic perovskite of BaRhO$_3$. (b) Orbital-resolved density of states. All the bands and density of states are plotted against the energy levels in the high symmetry reciprocal lattice of k-space. The electronic structure data are adapted from Materials Project database.[91](image)](image)

It is important to note that the orbital-resolved density of states, as shown in Fig. 1.3(b), contain crucial information about the adsorption properties of reaction molecules on the catalyst surfaces. According to the d-band theory,[135] the density of state distribution characteristics (e.g., center width, etc.) of adsorption site’s d-orbital plays an important role toward the degree of electron occupancy at anti-bonding molecular orbital, and thereby impacting the chemical bond strength. In the particular case of transition metal catalysts, an
1.4. Electronic factors in catalysis

increase of d-band center with respect to the Fermi level would decrease the electron filling of anti-bonding states, thus stabilizing the chemical bond. As the metal atom’s density of state contains rich information toward the catalyst surface reactivity, one of the primary tasks of machine learning study is to develop physically intuitive feature methods according to the DFT-calculated density of states. The following chapters provide detailed descriptions about the electronic structure features.
Chapter 2

Feature engineering of machine-learning models for alloy electrocatalysts

2.1 Introduction

Transition metals are versatile catalysts in directing molecular transformations. For a given metal, its surface reactivity can be tailored by varying geometric characteristics (e.g., coordination, strain), metal ligands, and extrinsic factors (e.g., solvation, support) in the vicinity of active sites.[186] With hierarchical complexities in catalyst design, a priori estimation of chemical reactivity of surface metal atoms is attractive. While all industrial catalysts used today were discovered via costly and time-consuming empirical testing, recent developments of density functional theory (DFT) have led to an unprecedented atomic-scale understanding of the complex processes occurring on catalyst surfaces and an identification of potentially improved catalysts from first-principles.[135] However, the immense phase space of catalytic materials spanned by structural and compositional degrees of freedom precludes thorough screening, even with combinatorial quantum-chemical calculations and high-throughput experiments.
2.1. Introduction

Herein we tackle this problem by developing a machine-learning chemisorption model that captures complex, non-linear adsorbate/substrate interactions through artificial neural networks (ANNs) and thus enables large-scale exploration of catalytic materials space. The proposed design strategy is built upon the kinetic analysis of surface reactions widely used to pinpoint underlying mechanisms and identify optimal adsorption properties of catalysts.[134] Within this design framework, adsorption energies of simple molecular fragments (e.g., *H, *CO, *OH, and *N) are reactivity descriptors extracted from the interrelated reaction network. Such simplification can be established because the chemisorption strength of a simple adsorbate on a given type of catalyst surfaces is linearly correlated with the adsorption energies of similar reaction intermediates including transition-state species,[6, 23] thus allowing us to directly build up relationships between reactivity descriptors and catalytic performance. While quantum-chemical density functional theory (DFT) provides a rigorous formalism[78, 101] to compute adsorption properties for a collection of hundreds of atoms with a reasonable accuracy, it is limited by the computational cost and system size. With the development of catalysis informatics, adsorbate/substrate interactions can be described using learning-based models without resorting to quantum-chemical calculations on a case-by-case basis.[11, 121, 127, 152, 193] Although the electronic structure characteristics, e.g., moments of the d-states distribution projected onto an active metal site,[194] are the compelling choice as input features in machine learning, the calculations of localized density of states requires convergence of electronic states at dense k-point sampling, which prohibits the on-the-fly prediction of catalytic materials. In this contribution, we go one step further to engineer numerical representation of surface atoms in machine-learning chemisorption models using easily accessible features, thereby allowing a facile prediction of surface reactivity of metal catalysts.
2.2 Computational methods

The density functional theory (DFT) calculations were performed using QUANTUM ESPRESSO\cite{57} with the Perdew-Burke-Ernzerhof (PBE) exchange-correlation functional.\cite{146} We used CO adsorption on the atop site of 100-terminated monometal and bimetallic surfaces with $4 \times 2 \times 4$ supercells and 15Å of vacuum space as model systems (core-shell ternary alloys is modelled using $4 \times 2 \times 5$ supercells). The geometric structures of model systems are shown in Fig. 2.1. The adsorbates and the two topmost metal layers were fully relaxed until the forces were smaller than 0.05 eV/Å while other metal layers were held fixed in their bulk positions. The Brillouin zones of all surfaces were sampled with $4 \times 8 \times 1$ Monkhorst-Pack k-points.\cite{125} The kinetic energy cutoff for plane-wave basis sets was 500 eV. The atomic cores were described by ultrasoft pseudopotentials.\cite{183} The occupation of Kohn-Sham eigenstates was smeared by the Fermi-Dirac function with width of 0.1 eV to aid SCF convergence for metallic systems, and all energies were extrapolated to $T = 0K$. The CO adsorption energy ($E_{CO}$) on a surface was calculated by,

$$\Delta E_{CO} = E_{CO/\text{slab}} - E_{\text{slab}} - E_{CO},$$

(2.1)

where $E_{CO/\text{slab}}$, $E_{\text{slab}}$, and $E_{CO}$ are the electronic energies of adsorbate/slab complex, reference slab, and the gas phase CO. The electronic energy of a gas phase CO was corrected by $-0.20$ eV due to the limitation of the PBE functional for describing gas phase molecules.\cite{29}

We use a feedforward artificial neural network (ANN) algorithm implemented in the open-source PyBrain\cite{161} code for a non-linear mapping between the input features and output target function, i.e., the adsorption energies. The details about the network configuration and input features will be discussed in the following sections.
Figure 2.1: Investigated configurations of \{100\}-terminated alloy models. (a) \(B@A_{ML}\), (b) \(A - B@A_{ML}\), (c) \(A_3B@A_{ML}\), and (d) \(Cu_3B - A@Cu_{ML}\). The first and second rows are top and side views of the structures, respectively. Group VIII and IB metals (Cu, Ag, Au, Ni, Pd, and Pt) were selected for both element A and B in the \(B@A_{ML}\) structure. In all other model structures, the d-block and post transition metals were considered for B, and A is chosen from the group VIII and IB metals.

2.3 Results and discussion

2.3.1 General strategy for machine-learning accelerated catalyst design

The schematic of the machine-learning accelerated catalyst design approach is shown in Fig. 2.2. The workflow starts with the data mining of ab initio adsorption energies from materials databases and standardizes the numerical representation of surface active sites using the geometry and composition of the systems. The datasets are then fed into an artificial neural network to build a connection between adsorption properties of catalytic sites and their numerical fingerprints. Once the network learned the underlying correlations of the system, the model can be used to screen through large materials space and suggest materials that
give favorable locations on the activity map obtained from the kinetic analysis. Interesting candidates will be evaluated from first-principles before experimental evaluations and the calculated data will be deposited into the database to further improve the model prediction and provide a feedback mechanism of the design cycle.

Figure 2.2: Schematic representation of the machine-learning accelerated catalyst design approach.
2.3. Results and discussion

2.3.2 Numerical fingerprints of metal surfaces

Feature engineering in machine learning is the process of designing numerical fingerprints of interested systems based on the domain knowledge. Identifying appropriate input features is the most fundamental and challenging step for the application of machine-learning methods. Many different types of features have been proposed to describe the thermodynamic and dynamic properties of materials.\[19, 44, 67, 68, 142, 166]\] For instance, the coulomb matrix, based on the Cartesian coordinates and nuclear charges of atoms, has been used as features for describing various thermodynamic properties of molecules and solids.\[44, 67, 68]\] The Behler-type symmetry functions are implemented in the high-dimensional neural network potentials for the on-the-fly prediction of interatomic forces in molecular dynamics simulations.\[19, 142]\] The partial radial distribution function (PRDF), which considers the distribution of pairwise interactions between two atom types, is used as fingerprints for electronic properties of crystals.\[166]\]

Development of physically transparent features that can capture adsorption properties of surface atoms is a key step to the machine-learning accelerated catalyst design. In Fig. 2.3, we illustrate available features based on our knowledge of the band theory of chemisorption. Among them, the moment characteristics\[194]\] of density of states projected onto the valence d-orbitals of a transition-metal atom, e.g., filling, center, width, skewness, and kurtosis, are important factors for understanding reactivity trends of metal surfaces.\[64, 100, 192]\] It has also been shown that the adsorbate-substrate bond length is strongly depended by the sp-electron density of substrate.\[192]\] Driven by the difference in electronegativity, the electron of metal alloys is transferred between a surface metal atom and its surroundings. To quantify the driving force of electron transfer in metal alloys, we define the local electronegativity of an adsorption site i (\(\chi_i\)) as the geometric mean of the Pauling electronegativity of metal atoms within the first neighboring shell.\[121, 193]\]
Figure 2.3: Feature engineering in machine-learning chemisorption models for metal catalysis.

\[ \chi_i = \prod_{j=1}^{1+nn} \chi_j^{\frac{1}{N}} \]  

(2.2)

where \( \chi_j \) is the electronegativity of atom \( j \) and \( N \) is the total number of atoms within the first neighboring shell including the adsorption site \( i \). We consider the local electronegativity as both geometry and electronic structure based feature.

To more efficiently describe the local chemical environment of an adsorption site, we propose the effective coordination number of an adsorption site as another geometry-based feature. The effective coordination number of an adsorption site is not a well-defined physical property. A simple bond counting scheme up to \( 2^{\text{nd}} \) nearest neighbors has been used recently
2.3. Results and discussion

for understanding reactivity trends of pure metal nanoparticles.[192] It has been shown that
the effective coordination of under-coordinated surface atoms defined using the accumulated
atomic charge density at the site normalized to that of isolated atoms linearly correlates
with the core-level shifts.[21] Inspired by those studies, we propose the effective coordination
number quantified by the interatomic d-d coupling matrix elements of an adsorption site
with its neighboring atoms up to the second nearest neighbors, defined as,

$$\text{CN}_{d}^{i} = \prod_{j=1}^{2^{nd_{nn}}} \frac{V_{dd}^{ij}}{V_{dd}^{i,\infty}},$$

where the interatomic d-d coupling matrix element[69] follows, $V_{dd} \propto (r_{d_{i}}^{3/2} r_{d_{j}}^{3/2} / d_{ij}^{5})$, $r_{d}$ is the
spatial extent of the metal atom’s corresponding d-orbital, $d_{ij}$ is the distance between the
atom i and j in the substrate, and $V_{dd}^{i,\infty}$ is the reference interatomic coupling matrix of pure
bulk of the host metal with the fully optimized geometry. The d-d interatomic coupling
is chosen because the d-band characteristics of a transition metal surface site is largely
determining the trend of surface reactivity based on the d-band chemisorption theory.[85]
In this study, we used the DFT optimized bond distances for computing the interatomic
coupling matrix elements. We note that various classical interatomic potentials, such as the
Embedded-Atom Method (EAM) potentials,[37] the bond-order potentials,[147] the Finnis-
Sinclair potentials,[178] and the effective-medium potentials,[89] can be used to optimize the
geometry of large metallic systems very efficiently. We introduced several secondary features
including the ionic potential, electron affinity, and Pauli electronegativity, which are only
host-metal dependent and can be fetched from literature or the periodic table. Inclusion
of those environment-independent properties is necessary for developing models involving
various different metals at adsorption sites.
2.3.3 Neural network configuration

We set up a feedforward artificial neural network using the open-source PyBrain code.[161] The network consists of a number of layers (input, hidden, and output) and each layer includes a few neurons (nodes) as the processing units. During the training process, each neuron receives the sum of weighted input from neurons of the previous layer and processes the cumulative input through the activation functions. The sigmoid activation function is implemented for the neurons in the hidden layers and a linear function is used for the output layer. The nodes in the previous layer and the subsequent layers are connected by a series of weight parameters, which will be systematically adjusted using the back propagation algorithms in the trainer to minimize the mean squared error between the target function and model hypothesis. Because the numerical value of different fingerprints could vary in the order of magnitude, feature standardization (scaling of numerical features to have the center at 0 and the standard deviation of 1) is applied routinely to improve the convergence of gradient-based algorithms. 75% of the available dataset is used as the training set and the rest 25% of the dataset is used for testing the model. We aim to prohibit the ‘over fitting’ of the model, which generally occurs when the network structure becomes excessively complicated (over trained) and fails in making predictions of the unseen data. To determine the optimal network structure with the most suitable number of neurons, we use the built-in Cross Validator module in the PyBrain to do the k-fold cross validation for the training set (see Fig. 2.4(a)). Using the geometry-based primary features together with secondary features, the network configuration with two hidden layers of 8 and 2 neurons gives the smallest cross validation error (<0.01 eV) (see Fig. 2.4(b)).

We first compare the neural network performance using the geometric and electronic features together with the secondary features, respectively, as illustrated in Fig. 2.5. For the geometry-based input features, the network configuration with two hidden layers of 8 and
2.3. Results and discussion

Figure 2.4: Neural network model architecture determination. (a) K-fold cross-validation scheme, and (b) the cross-validation errors as a function of the neural network structure using the geometric features and secondary features.

2 neurons was used, while the two hidden layers of 5 and 2 neurons were identified to give the lowest cross-validation error for the electronic structure based primary features.[121] Repeated randomization of data selection and training is used to avoid sampling bias and overly optimistic error estimates. Fig. 2.5 shows that the geometry-based fingerprints as primary features give similar prediction error (0.12 eV) compared to that using the electronic features (0.13 eV). Considering the simplicity of the geometry-based features and easy access to the geometric structure of systems compared to the d-band moments, the performance of machine-learning models for describing trends of surface reactivity of bimetallic alloys is excellent. It also suggests that the purely geometry-based local electronegativity and coor-
dination number can be used as indicators of variations in the local chemical environment of an adsorption site in response to strain and ligand engineering.

Figure 2.5: Evaluation of machine learning model prediction with DFT calculations. (a) the geometry based, and (b) electronic structured based primary features for CO adsorption energy on metal surfaces. The deviation of the RMSE from 16 random sampling is < 0.02 eV

2.3.4 Machine-learning models for catalyst prediction

We used CO2 electroreduction on 100-terminated metal surfaces as a model system to evaluate the performance of the neural network model. Some recent studies on CO/CO2 electroreduction over Cu(100) suggest that the dimerization of *CO with sequential electron and proton transfers governs onset potentials in C2 pathways, while a concerted proton-electron transfer to *CO is the critical step in C1 pathways [18,40]. Using linear scaling relations between adsorption energies of reaction intermediates and the adsorption energy of *CO, we calculate the theoretical limiting potentials for key elementary steps in CO2 electroreduction along C1 and C2 pathways as a function of the CO adsorption energy, shown in Fig. 2.6(a). Volcano-like reactivity curves for C1 and C2 pathways are observed where Cu is
close to the top, rationalizing the experimental observation of superior activity of Cu among transition metals for CO2 electroreduction. Decreasing the binding energy of *CO (less negative) on a Cu surface results in a less negative limiting potential for CO dimerization step. Based on this analysis, the optimal catalysts should have desired *CO adsorption energy in the shaded region, i.e., 0–0.2 eV weaker than that on Cu(100).

To evaluate the applicability of the neural network model with geometric features for catalyst design, we have used the neural-network model trained with all available dataset of bimetallics for predicting CO adsorption energy on the \( \text{Cu}_3B - A@\text{Cu}_{ML} \). This type of alloy has a monolayer of Cu metal atoms deposited on a four-layer alloy surfaces (e.g., \( \text{Cu}_3B \)), and one layer of another guest metal A as a buffer between them. This type of alloy has shown great flexibility in design due to a potential utilization of both strain and ligand effects for tailoring the reactivity of transition metals, and has been extensively studied for many electrochemical reactions, such as oxygen reduction in PEM fuel cells [41,42]. As a simple test, we show model-predicted CO binding energies on a subset of this type of alloys (\( \text{Cu}_3B - Ni@\text{Cu}_{ML} \) and \( \text{Cu}_3B - Rh@\text{Cu}_{ML} \)) as a function of DFT calculations. Fig. 2.6(b) shows that the neural-network predicted CO adsorption energy agrees reasonably well with self-consistent DFT calculations of CO adsorption (averaged RMSE \( \sim0.1 \) eV for 16 repeated random sampling in the network training), and several alloys are suggested to have desired CO adsorption energies and thus lower overpotential for \( CO_2 \) reduction.
2.4 Conclusions

In summary, we develop machine-learning chemisorption models using easily accessible, geometric features including the local electronegativity and the effective coordination number of an adsorption site, together with the intrinsic properties of active metal atoms such as the ionic potential, electron affinity, and Pauling electronegativity. Compared with the electronic structure based features, the machine learning model using the new representation of metal surfaces gives favorable prediction error $\sim 0.1$ eV for CO adsorption energies. The approach is used for screening 100-terminated multimetallic alloys for CO$_2$ electroreduction and shows some promising candidates with lower overpotential. The extension to other types of catalysts, such as realistic multimetallic surfaces and nanoparticles, transition-metal
2.4. Conclusions

dichalcogenides, and metal oxides, is the subject of ongoing work.
Chapter 3

Machine-learning models of bimetallic electrocatalysts

3.1 Introduction

In this chapter, the detailed strategy of machine-learning approach is demonstrated for the high-throughput exploration of the bimetallic electro-catalysts using the methanol electro-oxidation as model reaction with the aid of the established descriptor-based kinetic analysis. In recent years, direct methanol fuel cells (DMFCs) have shown great promise as an energy conversion device with a higher volumetric energy density than hydrogen-based fuel cells. However, the commercialization of this technology is largely hindered by the high over-potentials of methanol electro-oxidation at the anode side even using the state-of-the-art \{111\}-terminated \textit{Pt}-based catalysts. The high over-potential is primarily attributed to the strong bonding strengths of $^*CO$ intermediates on the \textit{Pt} surface. As \textit{Pt} catalyst surface can form strong chemical bond with the $^*CO$ adsorbate, the catalyst surface are heavily poisoned by the adsorbed $^*CO$ intermediates and less available activation sites are exposed for activating the water molecule into the $^*OH$. Subsequently, the overall chemical conversion to the final product of $CO_2$ is significantly limited due to the scarcity of $^*OH$ intermediates. To tackle $^*CO$ poising issue, the previous studies have shown some degree of success in designing novel bimetallic catalyst by alloying the \textit{Pt} with \textit{Ru} for an improvement
3.1. Introduction

of MOR catalytic activity, however, both the Pt and Ru elements are expensive materials in the real-world applications. In this regard, we presented a machine-learning framework to extend the search capability of density function theory (DFT) for identifying the {111}-terminated bifunctional catalysts. The bifunctional alloys, [58, 59, 165] which are termed as the metallic catalysts with geometrically adjacent but catalytically distinct sites, is able to activate the competitive rate-limiting steps that involving the CO* and OH* for lower over-potentials and reduced precious metal loadings.

Figure 3.1: Schematic display of the advantage of alloy-based catalyst’ bi-functionality in catalysis process. Catalytic reaction on (a) monofunctional catalyst and (b) bifunctional catalyst
3.2 Computational methods

All density functional theory (DFT) calculations were performed using Quantum ESPRESSO\cite{57} with the Perdew-Burke-Ernzerh (PBE) exchange-correlation functional.\cite{146} We used \{111\}-terminated metal surfaces with four layers \(2 \times 2\) surface cells as the model systems. The adsorbate and two topmost metal layers were fully relaxed until the forces were smaller than 0.05 \(eV/\AA\) while other metal layers were held fixed in their bulk positions. The Brillouin zones of all surfaces were sampled with \(6 \times 6 \times 1\) Monkhorst-Pack \(k\)-points.\cite{126} The kinetic energy cutoff for plane-wave basis sets was 500 \(eV\). The atomic cores were described by ultrasoft pseudopotentials. The occupation of Kohn-Sham eigenstates was smeared by the Fermi-Dirac function with width of 0.1 \(eV\) to aid SCF convergence for metallic systems, and all energies were extrapolated to \(k_B T = 0\) \(eV\). The binding energy of the \(\ast CO\) and \(\ast OH\) was calculated relative to \(CO\), gas phase \(H_2O\) and \(H_2\). We corrected the electronic energy of \(CO\) gas phase by \(-0.2\) \(eV\) because of the underestimation of \(CO\) formation energy using the PBE functional. The zero-point energy and entropy contributions are considered for free formation energies.\cite{49,157} The long-ranged van der Waals interactions are likely playing an important role for the adsorption energies of weakly adsorbed and/or highly polarizable species. The variation of \(CO\ast\) and \(OH\ast\) adsorption energies across metal surfaces is less dependent on the dispersion force. Thus, it is not considered in the calculations.

The potential-dependence of reaction free energies in elementary steps involving proton-electron transfers was evaluated using the computational hydrogen electrode (CHE) approach.\cite{133} In this approach, a reversible hydrogen electrode (RHE) is used as a reference:

\[
\frac{1}{2}H_2(g) \rightarrow H^+(aq) + e^- ,
\]  

(3.1)

where the proton-electron pairs are in equilibrium with \(H_2\) at all values of \(pH\) and 1 atm of
3.3 Results and discussion

$H_2$, without potential bias. At finite external potential $U$ vs RHE, the chemical potential of a proton-electron pair would be shifted by $-eU$, so we have,

$$\mu(H^+ + e^-) = \frac{1}{2} \mu(H_2) - eU.$$  \hspace{1cm} (3.2)

The theoretical limiting potential of an electrochemical reaction is defined as the minimal potential when all the steps become downhill in free energy.

3.3 Results and discussion

3.3.1 Kinetic analysis

A catalyzed chemical reaction typically involves a sequence of elementary events, e.g., adsorption of reactants onto a catalyst, bond rearrangement of reaction intermediates, and finally desorption of products into a continuum phase.[33] Rational control of the formation and cleavage of chemical bonds by tailoring the geometry and composition of an active site for efficient chemical and/or energy transformations is the ultimate goal of catalytic science. The conversational catalyst design strategy starts with a detailed understanding of the potential energy (Gibbs free energy) diagram for a given chemical reaction mechanism on the specific catalyst surface. DFT can approximate the thermodynamic stabilities for all the reaction intermediates along a reaction coordinate at each elementary step. Then, the potential or rate determining step are identified and the corresponding energy that required to drive all the reaction elementary steps become energetic favorable are calculated as the 'activity descriptors'. The previous literatures’ studies[49, 157] suggest that there are two coexisting reaction pathways (e.g., direct and indirect mechanisms) for the electro-oxidation of methanol on metal surfaces. In the direct mechanism, methanol is oxidized all the way
to $CO_2$ without explicitly involving the $CO^*$ intermediate; however, the activity with this mechanism is extremely low especially for $Pt$ because of the $CO^*$ poisoning, which dramatically limits the number of available active sites. The indirect mechanism which proceeds through the $CO^*$ intermediate have better activity on metal surfaces if $CO^*$ removal is not limiting. Based on this analysis, we focus particularly on the indirect mechanism in our study.

However, the conventional exhaustive approach requires a large amount of $DFT$ calculations, which can hardly meet the fast-growing industrial demands for novel cost-effective catalysts. To simplify the reaction network, the established theory about the 'scaling relationships'\cite{6, 23, 26, 48, 184} are applied to the methanol electro-oxidation reaction mechanism, where the inherent linear correlations among the bonding strengths of $CO^*$ with carbon-bonding species and $OH^*$ with oxygen-bonding species on a wide variety of metallic surfaces are quantified to give rise to a volcano-shaped relationship between the catalytic activity/selectivity and the adsorption energies of reaction intermediates, termed descriptors. It is noteworthy that the catalytic activity descriptor is defined as the theoretical limiting potential of an electrochemical reaction, which is the minimal potential that is required to drive all the consecutive elementary steps become downhill in free energy. Using the scaling parameters along with the free energy correction table that is provided by the literatures,\cite{49, 157} a volcano-shaped activity plot (see Fig. 3.2) is created to map out the theoretical limiting potential of methanol electro-oxidation as a function of two reactivity descriptors, i.e., the free formation energy of $CO^*$ and $OH^*$. To minimize the systematic error due to variations in computational methods, e.g., $DFT$ functionals, we used the $Pt$(111) surface as the reference in the reactivity descriptors.

The pure transition and noble metal surfaces are overlaid on the kinetic map in Fig. 3.2, which shows that the top region of the activity volcano is mostly occupied by scarce and
expensive noble metals, e.g., Rh, Ru, Ir, Os, and Re. Based on mechanistic analysis, the optimal catalyst should be active for methanol decomposition and water dehydrogenation, while allowing the removal of reaction intermediates, e.g., CO* and OH*, from an active site. Among pure metals, Pt is used commercially for a trade off between cost and capabilities in activating both methanol and water. Ru is the well known most active monometallic catalyst for water dehydrogenation and is typically added to Pt to form bifunctional alloys with alleviated CO* poisoning.[51, 189] The kinetic analysis above also reveals that the metallic Co and Ni are promising catalysts among all the mono-functional transition metals with the smallest limiting potential 0.5 V; however, they suffer from oxidation and stability problems under reaction conditions.[15]

Figure 3.2: Theoretical limiting potential (V vs. RHE) of methanol electro-oxidation as a function of CO* and OH* free formation energies relative to Pt(111). Pure transition and noble metals are overlaid on the map.
3.3.2 Machine-learning model development

Feature engineering

The design of fingerprint features is the most critical step for the development of machine-learning models. Feature selection from complex materials properties requires domain knowledge. There are a few considerations in developing fingerprint descriptors for adsorption properties of metal surfaces: (1) the features have to be unique in representing electronic and geometric structures of an active site; (2) it must be easily computed or readily available from databases to enable rapid screening; and (3) most importantly, they should be physically intuitive to ensure model robustness and direct inference of chemical insights. With that in mind, we developed primary features that can be obtained from DFT calculations, mainly characterizing the continuum $sp$-states and localized $d$-states of an adsorption site.\cite{111,121} The $d$-band features are represented by the moment characteristics of density of states projected onto the valence $d$-orbitals of a metal site, e.g., filling, center, width, skewness, and kurtosis.\cite{64,100,124,194} To account for the contribution of the substrate $sp$-electron density to adsorbate/substrate interactions, we include a geometry-based primary feature termed the local electronegativity ($\chi_i$), which is defined as the geometric mean of the Pauling electronegativity of metal atoms within the first neighboring shell ($d_{ij} < 3.5 \text{\textgreek{A}}$) of an active site. Those primary features were designed since they are important factors governing reactivity trends of various metal surfaces based on the d-band chemisorption theory and its recent developments.\cite{6,28,64,100,120,124,194} Besides the primary features, we have secondary features in fingerprint descriptors of an adsorption site including the ionic potential, electron affinity, Pauli electronegativity, work function, atomic radius, $d$-orbital radius, and $d$-band center of bonding metal atoms, which represent the intrinsic properties of host-metal atoms at an adsorption site. Those secondary features can be easily fetched
3.3. Results and discussion

from the solid-state table and can help distinguish different host metal-based alloys. [5] In this study, we have employed a broad range of model alloys with geometric structures shown in Fig. 3.3. By varying the A-site and B-site to metal elements across the periodic table, we obtained electronic properties and corresponding binding energies of simple adsorbates \( CO^* \) and \( OH^* \), the two most commonly used reactivity descriptors in heterogeneous catalysis, over one thousand of alloy surfaces. They are serving as the dataset for machine-learning model development.

![Figure 3.3: {111}-terminated bimetallic surfaces for the machine-learning model development, including A@B\(_{ML}\), A/B@A\(_{ML}\), A\(_3\)B@A\(_{ML}\), and A\(_3\)B, where A and B represents metal elements cross the periodic table.](image)

**Data preprocessing**

Since various electronic properties are used as fingerprint descriptors, a large variation of numerical values is present in the dataset, which can lead to low training efficiency and high prediction bias for the gradient-based optimization algorithms.[8] To alleviate this numerical issue, we standardize the fingerprint descriptors in the data preprocessing stage using,

\[
x_{std}^{(i)} = \frac{x^{(i)} - \mu_x}{\sigma_x},
\]  

(3.3)
where $x^{(i)}$ is the feature vector $x$ of a particular sample $i$, $\mu_x$ and $\sigma_x$ denote the average and standard deviation of the feature $x$. After standardization, our features are centered at the mean value of 0 with the standard deviation of 1. Nevertheless, outliers could exist in the samples, which potentially cause the compromised accuracy of model prediction. The violin plot in Fig. 3.4(a) maps out the distribution of the standardized fingerprint descriptors. The samples with any primary feature values outside of the statistical significance fences of $\pm 3\sigma$ are considered as outliers. Overall, 61 of 1032 metal alloys are recognized as outliers and are eliminated from future analysis. All datasets used for the training/validation of machine-learning models along with outliers are provided as the supplementary materials. The correlations among primary features are established in the Pearson’s product moment, given by the Fig. 3.4(b). Noticeably, all $d$-band features are strongly correlated, particularly for the $d$-band filling, center, and width as elaborated from the rectangular band model.[99] The local electronegativity is not correlated with the $d$-band moments because it is mainly characterizing the $sp$-band properties.

![Figure 3.4: Exploratory analysis of training data. (a) The violin plot of standardized fingerprint descriptors. The samples with any primary features outside of the $\pm 3\sigma$ bounds are outliers. (b) The Pearson’s pairwise product moment correlation heat map among the six primary features for model alloys.](image-url)
Machine-learning algorithms

A feed-forward artificial neural network was constructed to learn underlying correlations of fingerprint descriptors and the adsorbate/metal interactions using the open-source Keras library.[1] The network consists of different types of layers (input, hidden, and output) with a series of neurons as processing units at each layer. The nodes in adjacent layers are connected by a set of weight parameters, which will be systematically adjusted using the back-propagation algorithm in the trainer to minimize the mean-squared error between the target property and model prediction. During the training process, each neuron receives the sum of weighted input from neurons of the previous layer including a constant bias and processes the cumulative input through an activation function. The sigmoid activation function is implemented for the neurons in the input and hidden layers and a linear function is used for the output layer.

The available dataset is divided into two subsets with 75% for training neural nets and the rest 25% for testing the model performance. To prevent the overfitting, the validation dataset (25% of the training dataset) is used internally for the early stopping of the back-propagation trainer. Initially, the validation error keeps decreasing with iterations. However, when the network begins to overfit the training data and fails in making accurate predictions of the unseen data, the validation error starts to increase and the training process will be terminated. The 10-fold cross-validation is performed to identify the optimal network structure, e.g., number of hidden layers and number of neurons at each layer, using the Keras library.[1] The optimal structures for *CO and *OH datasets are 4 by 2 and 5 by 2 neurons with the two hidden layers, respectively. To eliminate the sampling bias, all the reported values were calculated based on 16-repeated randomization of data selection and training. The probability density of error distribution in insets of Fig. 3.5 shows that the residuals between the model-predicted and DFT-calculated adsorption energies for both
*CO* and *OH* are normally distributed and thus there is no sampling bias in the model. The root-mean-squared errors between the model prediction and the *DFT* calculation of both the 75% training datasets and 25% test datasets are ~0.2 eV. We have compared the training/testing performance of various machine learning algorithms in Fig. 3.6. As we can see the linear regression algorithms with varying regularization terms, (e.g., Linear regression, Ridge regression, Lasso, and Elastic net), give almost identical prediction errors for *CO* adsorption energies, ~0.28 eV, which are significantly worse than other algorithms. ANNs exhibit similar performance compared to the Kernel ridge regression, Support vector machine, Gaussian process, and K-nearest neighbor. The Random forest algorithm appears to be over-fit, manifested by smallest training error with limited generalization. In the following, we will choose neural network models because of its potential for meeting the challenge of growing materials datasets.

![Figure 3.5: Neural network training performance evaluation. *DFT*-calculated (a) *CO* and (b) *OH* adsorption energies on a set of idealized bimetallic surfaces vs. prediction from the machine-learning model in training and test. Insets show probability density of prediction errors for the training and test of the neural network model. RMSE errors of the model prediction are the average over 16 repeated data randomization to avoid sampling bias.](image)
3.3. Results and discussion

3.3.3 Model prediction of bimetallic catalysts

To design metal catalysts with improved performance, we test the applicability of the machine-learning model for rapid screening of bimetallic alloy surfaces that could synergistically activate both methanol and H$_2$O while not poisoned by surface intermediates. The neural network model trained with the dataset of mono-functional alloy surfaces is used for predicting CO* and OH* adsorption energies on six new types of bimetallic alloys as shown in Fig. 3.7. The structure of those types of bimetallic {111} surfaces contains a mixed monolayer of A and B surface atoms at varying ratios and either three bottom layers of the A metal or one B metal layer on top of two layers of the A metal. We extensively explore the materials subspace by substituting the A site with 11 FCC/HCP metals and B site with 25 transition and noble metals across the periodic table. We first perform DFT calculations of clean surfaces to obtain the electronic structure features as the input into the machine learning model and predict the adsorption energies of CO* and OH* on inequivalent surface sites. The theoretical limiting potential for a given catalyst with mixed surface layers is then computed using CO* and OH* free formation energies for the most favorable adsorption configuration of CO* and OH* in the limit of low coverage, i.e., the adsorbates interaction is ignored.

In Fig. 3.7(a), we show that the machine-learning model can quickly screen through a broad range of chemical space and predict many bimetallic catalysts with reduced overpotentials than Pt. Noticeably, the Pt/Ru near surface alloys with surface segregated Ru are predicted to exhibit enhanced methanol electro-oxidation activity with relatively small theoretical limiting potential <0.5 V vs. RHE. The Pt/Ru bimetallic alloys have been extensively investigated and many experimental measurements show that by alloying Pt with the oxophilic metal Ru, the electrocatalytic activity of methanol electro-oxidation is improved notably because Ru assists to dissociate the adsorbed water molecules to produce
Figure 3.6: Comparison to other regression algorithms in training and test of *CO adsorption energies on model alloys. RMSE errors of the model prediction are the average over the same 16 repeated data randomization to avoid sampling bias.

$OH^*$ species and $OH^*$ species can further oxidize the poisoning $CO^*$ species in the adjacent site to facilitate the catalytic process.[50, 116, 173, 190, 201] Among all the predicted surfaces, the near surface alloys with 3$d$-metal additives (e.g., Fe, Co, and Ni) are particularly attractive with low overpotentials. Those results complement many experimental observations that the Pt/Fe and Pt/Co nanoparticles show comparative electrocatalytic activity to the commercially available Johnson-Matthey Pt-Ru catalyst.[35, 38, 154, 195] The stability of those near surface alloys under relevant electrochemical conditions is an important but challenging issue to address. To add an extra layer of alloy screening, in Fig. 3.7(b), we show the $DFT$-calculated formation energies of near surface alloys relative to constituent atoms in their (pseudo-)fcc bulk crystal structures.
3.3. Results and discussion

3.3.4 Validation of machine-learning models

DFT calculations of the adsorption energies of randomly selected 40 A@A$_2$B$_2$ alloys are also performed to validate with the machine-learning models. Fig. 3.8 suggests that the machine-learning model has a reasonable accuracy (~0.25 eV) compared to self-consistent DFT calculations. In comparison to previously developed linear models using similar electronic structure features, the machine-learning model [127, 193] developed here can be applied to a much broader alloy space for screening.

3.3.5 Feature importance analysis

The artificial neural network model achieved through a set of highly correlated weight parameters is complex and typically considered as the black box. To physically interpret the neural network model and elucidate the underlying factors that govern the strength of adsorbate/metal interactions, we perform sensitivity analysis based on the perturbation method [43]. The feature importance is measured using the normalized sensitivity coefficient (NSC$_i$) that captures the overall response of the network’s output targets, i.e., adsorption energies, to a 25% perturbation of a given input feature $f^i_j$:

$$NSC_i = \frac{1}{N} \sum_{j=1}^{N} \left| \frac{\partial \Delta E_j}{\partial f^i_j} \right| \cdot \frac{\max(f^i) - \min(f^i)}{\max(\Delta E) - \min(\Delta E)}$$  \hspace{1cm} (3.4)

where $f^i_j$ is the feature i of the sample j and $\Delta E_j$ represents the output of the sample j, i.e., adsorption energies of CO* or OH*. The larger value of the normalized sensitivity coefficient suggests a stronger dependency of the target property to a given input feature. Fig 3.9 shows the normalized sensitivity coefficients for both the CO* and OH* adsorption energies with respect to 6 primary features. The standard deviation of the NSC$_i$ is also shown for
the evaluation of 16 machine-learning models derived from random data sampling. There are multiple insights that can be drawn from the feature importance analysis: 1) the $d$-band features that are characterized by the lower moments (e.g., filling) play a more important role than higher moments; 2) $CO^*$ has a stronger dependence on $d$-band features than $OH^*$ because $CO$ has unoccupied $2\pi^*$ molecular orbitals right above the Fermi level for hybridization, while the $1\pi$ orbital of $OH^*$ is almost fully occupied and the $4\sigma^*$ orbital is further above the Fermi level;[65, 192] and 3) $sp$-band properties indirectly determined by the local electronegativity via charge transfer has a noticeable effect on $OH^*$ compared to $CO^*$ due to Pauli repulsion between interacting non-orthogonal orbitals, consistent with previous models.[192, 193] This insights will have important consequence by providing feasible strategies to break energy scalings of strongly correlated reaction intermediates in catalysis,[7] for example, $O^*$ vs. $OH^*$ in oxygen reduction, $CO^*$ vs. $CHO^*$ or $COH^*$ in $CO_2$ reduction, and $N^*$ vs. $NH_x^*$ in $N_2$ reduction.
3.4 Conclusions

To conclude, a holistic machine-learning design framework is developed to accelerate the discovery of bimetallic catalysts. This approach can rapidly explore a broad chemical space with enormous compositional and configurational degrees of freedom, and narrow down interesting candidates for certain applications prior to high-level quantum-chemical calculations and/or experimental testing. By learning from the data of ~1000 model alloys with their electronic and adsorption properties, the adsorption energies of $CO^*$ and $OH^*$, two commonly used reactivity descriptors in heterogeneous catalysis, can be predicted for new bimetallic catalysts with $RMSE \sim 0.2$ eV. To validate this approach, we have shown that the model successfully identifies most of known alloy catalysts for methanol electro-oxidation and suggests 3$d$-metals, e.g., Fe, Co, and Ni, as promising additives for bimetallic catalysts with reduced overpotentials. Feature analysis indicates that the improved activity of bimetallic surfaces and the justification of the two descriptors approach originates from distinct governing factors of $CO^*$ and $OH^*$ adsorption on metal surfaces.
Figure 3.7: High-throughput screening of novel bimetallic alloys enabled by neural network model. (a) Model-predicted theoretical limiting potentials of catalyst surfaces towards methanol electro-oxidation based on the machine-learning model predicted $CO^*$ and $OH^*$ free formation energies and kinetic mapping. The geometry structure of this type of bimetallic catalysts is shown on the top with varying ratio of $A$ and $B$. (b) $DFT$ calculated formation energy relative to atoms in their (pseudo-)fcc bulk crystal structures. The grey out squares represent systems that are reconstructed during structure optimization.
Figure 3.8: Comparison between model-predicted and DFT calculated adsorption energies for CO* and OH* on 40 randomly selected A@A₂B₂ type bimetallic catalysts. Pure Pt is shown for comparison.
Figure 3.9: Feature importance scores for $CO^*$ and $OH^*$ models based on the sensitivity analysis of neural network models. The inset shows the clustering and linear dependence among input features and target properties in the linear model.
Chapter 4

Machine learning energy gaps of metalloporphyrins

4.1 Introduction

The previous Chapter primary illustrates the predictive functionality of the supervised
machine-learning algorithms for designing the anode side catalytic material of the direct
methanol fuel cells (DMFCs). In Chapter 4, the interpretative functionality of machine-
learning technique is fully demonstrated by exploring the Porphyrin’s structure-property
relationship with the aid of a global sensitivity analysis towards the physically intuitive
structural descriptor (electrotopological-state index). More specifically, We will be focused
explicitly on identifying the underlying structural factors that governing the energy gaps
property using the machine-learning techniques and the insightful knowledge would be ex-
tremely valuable for explaining the porphyrin’s electron transfer effect that determined by
the inherent structural components. Fig 4.1 provides a conceptual description of the study.

Porphyrins, a family of macrocyclic compounds, contain four pyrrole rings joined by methine
bridges.[22, 77] Attributed to their highly versatile structures, the physicochemical properties
of porphyrins can be tailored via substituent ligands and chelating metal ions with theoreti-
cally unlimited combinations.[84, 114, 115, 160, 198] Arguably, the energy gap is one of the
most important characteristics for light-harvesting applications.[16, 140, 141] Porphyrins
with appropriately positioned energy levels of frontier orbitals have shown great promise in solar-to-electrical energy conversion,[10, 12, 70, 71, 72, 107, 118, 123, 139] photosensing,[20] photomedicine,[63, 81, 94, 149] photocatalysis,[32, 42, 46, 187, 199] etc. Traditionally, the search for such materials through experimental trial and error is extremely tedious, costly, and time consuming. Recent advances in computational infrastructures and electronic structure methods, e.g., density functional theory (DFT), allow us to design materials from first principles. However, the immense materials space spanned by a vast number of substituent ligands and chelating metal ions of porphyrins prohibits high-throughput screening due to the high computational cost of quantum-chemical simulations with accurate exchange-correlation functionals.

In recent years, there has been a dramatic rise in the employment of machine-learning techniques for understanding materials chemistry and accelerating materials discovery attributed to the increasing accessibility of large-scale, open datasets of materials and high-performance computing infrastructures.[14, 24, 39, 92, 106, 111, 112, 121, 138, 150] In contrast to quantum-chemical methods that solve the Schrödinger equation of a collection of atoms from first principles, supervised machine learning relies on labeled data to build a mapping of feature representations onto properties of interest using regression or classification techniques. Numerically representing a material or its local structural information in terms of relevant physical factors is crucial for machine-learning model development. In this regard, numerous efforts have been made. For instance, the electronic and geometric structure based features including moments of the d-band distribution of transition metals,[111, 112, 121] the molecular representations of heterogeneous set of molecules, [45] atomic neighbor environments descriptor,[17] the Gaussians described molecular one-body potentials, [25] atomic environment vector that described by a modified Behler and Parrinello symmetry functions, [171] the DFT-calculated bonding strengths of the frozen atomic system, [53, 170]
4.1. Introduction

the crystal’s generalized constituent element properties, [108, 148] the partial radial distribution functions,[90, 166] Coulomb matrix,[159] motif based fingerprints,[80] topologies of crystal structures,[47] Morton space-filling curves,[93] and electronic structure diagrams,[86] have been utilized to directly correlate with materials properties or performance indicators. Physical insights can be drawn for human learning, although it is very challenging and not always pursued by machine learning practitioners and data scientists. In this work, we develop highly accurate machine-learning models for predicting energy gaps of porphyrins with molecular graph representations and extract physically interpretable insights into governing structural motifs with feature sensitivity analysis and quantum-chemical calculations.

Figure 4.1: Conceptual visualization of Porphyrin’s structure-property relationship enabled my machine-learning approach.
4.2 Computational methods

4.2.1 DFT specification

All DFT calculations were performed by QUICKSTEP in CP2K package.\cite{82} The exchange-correlation potential was treated with the generalized gradient approximation parameterized by the spin-polarized Perdew-Burke-Ernzerh functional.\cite{146} The wavefunctions were expanded in molecular optimized triple-$\zeta$ valence polarized Gaussian basis sets (TZVPMOLOPT-GTH), with an auxiliary plane-wave basis set with a cutoff energy of 320 Rydberg. Core electrons have been modeled by the scalar relativistic norm-conserving Goedecker-Teter-Hutter (GTH) potentials with 4, 1, 5, 6 and 6 valence electrons for C, H, N, S and O, respectively. In the geometry optimizations, the maximum forces were converged to 0.1 eV/Å.

4.2.2 Dataset description

The dataset for machine learning was retrieved from the Computational Materials Repository,\cite{105} which includes >12,000 molecular computational structures of porphyrins and DFT-calculated properties, e.g., frontier orbital energy levels, optical gaps, and energy gaps. In this study, we focus on the energy gap calculated as the difference between the electron affinity and ionization potential, i.e., $E_{ea} - E_{ip}$, because of its importance in determining efficiencies of solar light absorption and energy transfer.\cite{10, 12, 70, 71, 72, 107, 118, 123, 139} By varying the side groups $R_1$, $R_2$, $R_3$, and the anchor group $R_4$ at meso-positions, peripheral substituents $L$ at $\beta$-positions, and chelating metal ions $M$ as denoted in Fig. 4.2, a theoretically unlimited number of porphyrins can be conceived and synthesized. In the dataset, the side groups are aromatic ligands and the anchor group is connected to the methine bridge via a
carbon-carbon double or triple bond and a carboxylate group as the anchor point to semi-conducting supports, e.g, TiO$_2$.\cite{123, 139} For the porphyrins in the database, $R_1$ and $R_3$ ligands are kept the same, while $R_4$ has two rotational configurations with respect to the porphyrin plane. A complete list of functional groups used in the dataset can be found in the Computational Materials Repository.\cite{105}

Figure 4.2: Structural labeling scheme of porphyrins with varying ligand substitution and metal chelation.
4.3 Results and discussion

4.3.1 Machine-learning model development

Molecular graph representations

Molecular representation, or broadly speaking feature engineering, is a key ingredient in the development of machine-learning models. The process numerically transforms the valuable information encoded in a molecular structure into a high-dimensional vector that is invariant with respect to rotations, translations and ideally has as minimal redundant information as possible. As shown in Fig. 4.3, four different types of molecular graph representations, i.e., molecular fingerprint, Coulomb matrix, chemoinformatics, and electrotopological-state index, are used in our study. The molecular fingerprint enumerates the fragment components in the molecular graph and maps the fragments onto a 1024-dimensional binary bit representation, where each bit indicates the presence (1) or absence (0) of a specific subgraph.[30, 130, 153] The OpenBabel library is used for generating molecular fingerprints of porphyrins.[137] The Coulomb matrix represents a 3D molecular structure in a 2D array tabulating the strength of Coulombic interactions between atoms according to their Cartesian coordinates and nuclear charges,[159] see Eq. 4.1,

\[
C_{ij} = \begin{cases} 
0.5Z_i^2 & \forall i = j \\
\frac{Z_i Z_j}{|R_i - R_j|} & \forall i \neq j.
\end{cases}
\] (4.1)

The sorted eigenvalues (eigenspectrum) of the Coulomb matrix or its truncated versions[67] have been used as the feature of a given molecular structure. The cheminformatics vectorizes the molecular graph’s global physicochemical properties defined by its structural, electronic, and quantum-chemical characteristics.[129] With the package Mordred,[2, 129]
we retrieved a set of 13-dimensional cheminformatics features including the molecular connectivity, fragment complexity,[132] rotation bonds ratio, electronic polarizability, etc. The electrotopological-state index (E-state index) has been widely used in Quantitative Structure-Property Relationship (QSPR) models to predict boiling points, aqueous solubilities and other thermophysical properties of organic molecules.[34, 61, 88, 131, 158] It considers a molecular graph as a correlated information field.[97] Within a given chemical environment, an atom’s intrinsic electrotopological state, $I_i$, is computed by considering its local chemical bonding. An intuitive explanation of $I_i$ is the valence-state electronegativity[41] defined by Eq. 4.2:

$$I_i = [(2/n)^2\sigma' + 1]/\sigma \quad (4.2)$$

where $n$ represents the atom’s principle quantum number of valence electrons, $\sigma$ and $\sigma'$ denotes the number of all-type electrons and valence electrons participating sigma bonding, respectively. The E-state index for a given atom in the molecule is influenced by all the surrounding atoms, with the perturbation $\Delta I_i = \sum_{j=1,j\neq i}^N (I_i - I_j)/r_{ij}^2$, where $r_{ij}$ denotes the number of chemical bonds between any two atoms in their shortest graph pathway and $N$ represents the total number of atoms in a molecular structure. The E-state index of the atom $i$, $S_i$, is then computed as the sum of the intrinsic E-state value with the perturbation, $S_i = I_i + \Delta I_i$. The atom’s E-state index encodes the unification of electronic and topological attribute information within a given chemical environment, and it is analogous to electronegativity for conceptual understanding.[97] A lot of investigators believe that the atom’s E-state index value is strongly correlated with the intermolecular interactions, e.g., electrostatic and hydrogen bonding strength.[41, 62, 97] To extract the E-state indices of a given porphyrin, the entire information field is decomposed into a 15-dimensional labeled atom centroids according to their connectivity with neighboring atoms.[96] For instance, the $dssC$ denotes the carbon motif that is connected to the surrounding atoms with one double
bond and two single bonds. Subsequently, the sum of all the $E$-state index values for the same-type atoms in a molecule is used as one component of the feature. To appropriately describe all the porphyrins in the dataset, we added a 9-dimensional categorical vector to denote the presence (1) or absence (0) of a particular chelating metal ion and anchor group rotation in the structure.

![Diagram of molecular graph representations](image)

Figure 4.3: Schematics of molecular graph representations: (a) molecular fingerprint, (b) Coulomb matrix, (c) cheminformatics, and (d) electrotopological-state index.

**Machine-learning algorithms**

The performance of the four molecular graph representations in predicting the energy gaps of porphyrins is evaluated by a variety of machine-learning algorithms, e.g., Lasso, kernel ridge regression (KRR), support vector machine (SVM), and feed-forward artificial neural networks (ANNs) with sigmoid activation function for hidden layers and linear function on
4.3. Results and discussion

The dataset is randomly divided into 75% for training/validation and the rest 25% for test. To improve the algorithm training efficiency and interoperability, all the input features are scaled with 0 mean and the standard deviation of 1. In this work, we use the Keras library[3] for the feed-forward ANNs and the scikit-learn library for other machine-learning algorithms.[52] The 5-fold cross-validation technique is used to determine the optimal model parameters for the machine-learning models, i.e., the neural network architecture (2 hidden layers with 10 neurons on each layer), the regularization parameter \((\lambda = 0.0003)\) of Lasso model, and the RBF kernel coefficient \((\gamma_{KRR} = \gamma_{SVM} = 0.01)\) and the regularization terms \((\alpha_{KRR} = 0.001, C_{SVM} = 10)\) for both the kernel ridge regression and support vector machine.

**Sobol sensitivity analysis**

Sobol sensitivity analysis is a probabilistic approach that aims to quantify the contributions of a single input variable \(X_i\) or a set of interrelated variables \((X_1, ..., X_p)\) to the overall model output unconditional variance \(Var(Y)^3\). Under the assumption that the input variables are independent, Sobol theory suggests any model function \(f\) can be decomposed as the sum of partial functions \(f_{i,...,p}(X_i, ..., X_p)\) with a total \(p\)-dimensional feature space in the following way:

\[
Y = f(X_i, ..., X_p) = f_0 + \sum_{i=1}^{p} f_i(X_i) + \sum_{i=1}^{p} \sum_{j=i+1}^{p} f_{ij}(X_i, X_j) + ... + f_{i,...,p}(X_1, ..., X_p),
\]

where \(f_0\) is the expectation value of model output \(E(Y)\). Assuming that the model function \(f\) is square-integrable, as an example, the unconditional variance that associated with the partial function of \(f_{ij}(X_i, X_j)\) in Eq. 4.3 is given by, \(Var_{i,j} = \int_0^1 \int_0^1 f^2(X_i, X_j) dX_i dX_j\). In this way, the total model unconditional variance \(Var(Y)\) is successfully decomposed into a set of
variances that are related to the partial functions. The contribution of individual input’s variance to the total output variance can be determined by the ratio of the partial variance to the total variance, the Sobol sensitivity indices. The Sobol sensitivity indices represent the fractions of the unconditional model output variance that are associated with the input variables. In reality, the variance of each partial function is a complicated integration term, which requires extensive numerical effort to solve. However the Monte Carlo sampling approach efficiently solve those integrations by inference.

### 4.3.2 Model prediction of energy gaps

The training/test parity plots are shown in Fig. 4.4 for different features and learning algorithms. An average of RMSE errors from 16 runs is presented as one of the performance indicators, while the uncertainty of model prediction is computed by the bootstrap sampling approach\[74\] to provide a full description of the model performance. The bootstrap sampling method efficiently assesses the model statistical accuracy and uncertainty through a repeated data random draw and model refitting process. The insets in Fig. 4.4 show the histogram of the probability distribution of the bootstrapping errors with 100 measurements. The red vertical line marks the mean of model prediction errors and the shaded area represents the 95% likelihood confidence interval. The normality of the distribution of bootstrapped errors reveals the model prediction stability. After taking into account all the above criteria including the RMSE, bootstrapped mean, and 95% bootstrapped confidence intervals, the $E$-state index is identified as the best molecular graph representation for predicting the energy gaps of porphyrins. For this representation, KRR, SVM, and the feed-forward ANNs give similar model performance with RMSE $<0.06$ eV, while the Lasso method gave a wide-band parity plot, indicating poor model performance.
The model training and prediction results in Fig. 4.4 assess the performance of the molecular graph representations from the perspective of the error distributions. Another important perspective is their generalizability to capture a broader chemical space. In this regard, it is necessary to address the representation methods’ advantages and disadvantages in terms of the inherent data structure from different methodologies. As one of the advantages, all the representations can be efficiently generated from the molecular graphs without relying on any sophisticated electronic structure optimization. The drawback of the molecular fingerprint is the high-dimensional (1024) feature space and the lack of sensitivity (bias) to the small fluctuations of molecular structures. The high dimensionality would cause significant increase of the computational costs of machine-learning algorithms especially with kernel functions, i.e., KRR, SVM, and the bad sensitivity would allow the method failed to differentiate any molecular compounds with the same structural components. Similar to the molecular fingerprint, the Eigenspectrum of Coulomb matrix, regardless of the high dimensionality, also has high bias to capture the underlying structure-property relationships because of the information loss in the Eigen-decomposition of a molecular Coulomb matrix. In contrast, the lower dimensional cheminformatics and $E$-state index, have excellent bias–variance tradeoff for the small and almost identical training and prediction errors as shown in Fig. 4.4. The $E$-state index is more attractive than the cheminformatics not only because of the relative lower prediction error but also its interpretability.
Figure 4.4: Parity plots for DFT-calculated vs. machine-learning model prediction of energy gaps with molecular graph representations. The insets show the probability distribution of bootstrapping errors. The shaded region represents a 95% likelihood confidence interval and the red vertical line denotes the mean of bootstrap errors.
4.3. Results and discussion

4.3.3 Principal component analysis

To rationalize the superior performance of the $E$-state index as a feature, we performed Principal Component Analysis ($PCA$) of the dataset for all molecular graph representations. In essence, the PCA projects a high-dimensional feature onto a visible 2-dimensional subspace, where the variance of a dataset is maximized.[144] By visualizing the relative locations of the samples on the PCA-compressed subspace, we can obtain a general sense of the molecular descriptors’ similarity. As shown in Fig. 4.5, the scatters of available samples, characterized by two principal components, densely overlap for the Coulomb matrix and molecular fingerprint representations, while they are nicely distributed for the $E$-state index and to a lesser extent for the cheminformatics. This analysis partially rationalizes the trend of model performance using various features shown in Fig. 4.5 and can be used as a guide for feature selection.

4.3.4 Model interpretation

Another advantage of the $E$-state index is its capability to encode the molecular structure with atom-level information, thus allowing the assignment of the electrotopological property for a particular structural motif. This crucial characteristic enables the $E$-state index to uncover the insightful knowledge about the structure-property relationships of materials. Historically, machine-learning models are “black-box” mappings of the sophisticated intercorrelations of output with features, in this case between the energy gaps of porphyrins and molecular graph features. It is extremely challenging to measure the contribution of an individual descriptor to the property. To tackle this challenge, we employed a global variance-based sensitivity analysis approach introduced by Sobol[172] to qualitatively determine the most influential structural attribute with respect to the energy gaps. This analysis
has been very useful for model builders to interpret complicated models in the applications such as biological systems,[98] pharmacology,[200] and risk assessments.[181] In essence, the Sobol sensitivity analysis is a probabilistic approach that aims to quantify the contributions of each input variable to the unconditional variance of model output.[172] The Sobol sensitivity indices represent the fractions of the output variance associated with each input feature using Monte Carlo sampling and bootstrapping implemented in the SALib package.[75]

The polar plot in Fig. 4.6(a) shows the feature importance of the E-state index using various machine-learning algorithms. For ease of visualization, only influential features with indices larger than 0.01 are plotted. Fig. 4.6(a) indicates that the E-state index of the pyrrole ring carbon that is connected to the anchor group with a single bond (aasC) is the
most influential structural motif for the energy gaps of porphyrins, irrespective of learning algorithms. To understand the observed trend of feature importance, we have performed DFT calculations of a pristine porphyrin (no metal chelating) with Ethen-ThioPhene-Cyano-Acrylic as the anchor group and \( H \) atoms at all side groups. As seen in Fig. 4.6 (b), density functional theory calculations show that the \( HOMO \) state is delocalized over the pyrrole ring, while the \( LUMO \) state is localized on the anchor ligand, facilitating the excited electron transfer into the conduction band of oxide supports, for example in dye-sensitized solar cells (DSSCs).[10, 12, 70, 71, 72, 107, 118, 123, 139] Since the \( aasC \) motif has a larger spatial overlap with the \( LUMO \) state of the porphyrin while the \( HOMO \) state is delocalized over different types of \( C \) atoms in the pyrrole ring, the electronic and topological properties of this atom motif strongly correlates with the energy gaps.

Figure 4.6: Interpretation of machine learning models using Global sensitivity analysis. (a) Polar plots of Sobol sensitivity indices for visualizing the influence of key structural attributes with respect to the energy gaps of porphyrins. In the molecular graph theory, the convention of labeling is \( s \): single bond, \( d \): double bond, \( t \): triple bond, and \( a \): aromatic. (b) Isosurface of the calculated frontier molecular orbitals of a typical porphyrin with the Ethen-ThioPhene-Cyano-Acrylic as the anchor group.
4.4 Conclusions

In conclusion, we have developed highly accurate machine-learning models for predicting the energy gaps of porphyrins, an important property for light-harvesting applications. Among various molecular graph representations, the $E$-state index shows a robust prediction RMSE $<0.06$ eV with various learning algorithms, e.g., artificial neural networks. The optimized machine-learning model can be further applied for exploring high performance porphyrins once the molecular graph representations are extracted from a novel computer-synthesized porphyrin structure. A global variance-based sensitivity analysis suggests that the carbon structural motif on the methine bridge connected to the anchor group contributes the most in modulating the energy gaps, consistent with the spatial distribution of their frontier molecular orbitals from quantum-chemical calculations. The machine learning models developed here have potential to facilitate the design of porphyrins with appropriately aligned oxidation potentials relative to a typical $I^-/I_3^-$ redox couple and the conduction band of semiconductors in DSSCs, and other relevant types of materials, e.g., metal oxides, molecular complexes, and metal-organic frameworks. A full machine-learning analysis of existing datasets in chemical science community could provide physically interpretable models for guiding materials discovery, but more importantly, for human learning.
Chapter 5

Adaptive machine learning for perovskite electrocatalyst design

5.1 Introduction

Recent years have witnessed an emergence of artificial intelligence in high-throughput materials design.\cite{151, 162, 164} Within this domain, machine learning algorithms offer the promise of establishing the correlations between material representations and the properties of interest. To date, machine learning has become one of the most fascinating tools in modern materials research and a number of interesting applications have emerged. For instance, advanced machine learning algorithms are applied to predict the first-principles calculated properties such as formation energies of crystals,\cite{40, 44, 56, 163, 191} metal alloys’ adsorption properties,\cite{111, 112, 121, 180, 182} electronic properties,\cite{31, 166} and optical characteristics of metal oxides and metalloorganics.\cite{113, 119, 148, 188} Besides, machine learning algorithms are also utilized to estimate inorganic materials’ physicochemical properties such as melting temperatures,\cite{168} heat capacity,\cite{95} and optical properties.\cite{174} Typically, those bulk material data are available either in the material databases or can be generated from high-throughput computational infrastructures. While in many cases, collecting a sufficient amount of material data for model training is not always feasible, particularly for catalysts, in which reactions occur at surfaces, defects, and boundaries. One of the primary
reasons is that the description of the surface reactivity for those systems requires accurate quantum-chemical methods. Consequently, the machine learning models optimized with small training data tends to be overfitted and could provide unreliable estimates due to the unbalanced bias-variance tradeoff. Owing to the inherent capability of uncertainty assessment, probabilistic models (e.g., Bayesian linear regression, Gaussian process regression) have received enormous attention in the fields of medical diagnosis, financial forecasting, robotics technology, etc. In the meanwhile, uncertainty estimation is becoming a crucial aspect in computational materials design and machine learning. Many material studies are reported with a particular focus on uncertainty quantification. For instance, the adaptive learning schemes enabled by probabilistic models are used to estimate the morphotropic phase boundary of piezoelectric crystals, the electronic band structures of hetero-structures, and the thermal hysteresis of NiTi-based shaped memory alloys.

In this work, we address the value of uncertainty in guiding the machine learning model development with a case study of designing double perovskite catalysts for the oxygen evolution reaction (OER), which is the bottleneck in many energy capture and storage devices. Perovskites with a ABO3-type structural configuration have received great attention for the environmental and energy-related applications attributed to their superior electrical, magnetic, optical, and catalytic properties. The composition of perovskites is versatile with a wide variety of element selection for the A and B sites, thus leading to a theoretically infinite materials space. Typically, the A-site cations are rare-earth or alkaline-earth metals, while the B-site cations are transition metals. Even though enormous efforts have been devoted experimentally to explore inexpensive perovskites with comparable activities to the state-of-the-art (iridium/ruthenium oxide) catalysts, the search capability with the trial-and-error approach is limited. With recent advances in high-throughput experimentation and computing infrastructures, the descriptor-based approach
5.1. Introduction

has gained popularity in catalyst screening. Reactivity descriptors, including the number of d-electrons of B-site cations, oxygen p-band center, O-M covalency and the free energies of oxygen-containing species (e.g., O* and OH*), have been explored to characterize the catalysts’ intrinsic activities toward water oxidation. We are particularly interested in the structure-activity relationships enabled by the adsorption-energy scaling relations of reaction intermediates. The framework greatly simplifies the multi-step reaction network to a large extent and has shown promising results in guiding catalyst discovery. However, it is time-consuming to compute the reactivity descriptors using the electronic structure methods due to the strongly-correlated electrons of transition metal ions and the infinite materials space, which calls for a new approach.

We demonstrate that the uncertainty-guided machine learning is an effective strategy to tackle the challenge of data deficiency for catalyst design. The approach is dependent on physically intuitive material representations and an automatic, adaptive learning process enabled by probabilistic modeling. In this work, we use a heterogeneous feature set with multi-fidelity informatics, i.e., bulk electronic structure and atomic properties. Gaussian process regression models were developed to estimate the Gibbs free energies of key reaction intermediates, e.g., O* and OH* at B-site cations. The probabilistic Gaussian process is chosen for the assessment of model uncertainties. The uncertainty is used as a guide to progressively improve the model prediction in a closed loop where the expensive DFT calculations are selectively performed on the model-predicted candidates with relatively small overpotentials (<0.5 V) albeit with large uncertainties. In this strategy, it is not necessary to develop a fully optimized model for the entire material space. After the refinement process, the ratio of the reliable candidate dataset is improved by ~40% compared to the initial model with randomly-picked 250 samples. Importantly, attributed to the physically intuitive descriptors, we draw insightful knowledge that the $e_g$-orbital of B-site cations plays a crucial
role in OER by interpreting the feature distributions of better and worse candidates than LaCoO$_3$.

Figure 5.1: Schematic of the adaptive machine-learning workflow.

5.2 Computational methods

DFT is performed to calculate the adsorption energies of \( ^*\text{O} \) and \( ^*\text{OH} \) at the B site of perovskites along with the bulk electronic structures. The Vienna Ab initio Simulation Package (VASP) with the revised Perdew–Burke–Ernzerhof (RPBE) generalized gradient approximation (GGA) exchange-correlation functional[66] is chosen in consideration of a balance between accuracy and computational cost. The kinetic energy cutoff for the plane-wave basis set is 500 eV. The 001-terminated cubic (Pm3m) $ABO_3$ perovskites with a $2\times2\times4$ super cell is constructed as the model system. The Monkhorst–Pack k-points grid was set up with $4\times4\times1$. The adsorbates ($O^*$ and $OH^*$) are attached at the atop of the B cation at the -$BO_2$ termination and three topmost model layers were fully relaxed while other layers
were held fixed in their bulk positions during the DFT optimization. It is noteworthy that the GGA functional couldn’t capture the strong on-site Coulomb interactions for specific d-electrons, therefore we employ the standard Hubbard U approach from Materials Project\cite{76} to the B-site elements (e.g., Mo, Ni, V, Co, Fe, W, Mn, Cr) for a better description of the underlying electronic structure. Besides, the spin-polarization calculations are carried out for the perovskite structures containing the B-site elements of Cr, Mn, Ni and Co.

5.3 Results and discussion

5.3.1 Adaptive learning strategy

Essentially, adaptive learning attempts to improve the reliability of machine learning models with an iterative training/validation process. During the adaptive process, the model’s initial weak hypothesis toward the materials’ structure-property relationship is progressively adjusted by adopting the evidence. In this work, the model’s weak assumption is primarily caused by the insufficient training data. As a consequence, the model tends to be overfitted and is likely to make significant mistakes toward the properties of unseen materials. To tackle this challenge, we select candidates for DFT validation by following the rule of balancing the tradeoff between exploration and exploitation.\cite{151} That is, special attention should be paid to those candidates, which have the promising catalytic activity albeit with a high uncertainty. Under the guidance of uncertainties quantified by the probabilistic mode, adaptive learning procedure allows us to gradually improve the model performance with only a small computational effort. Fig. 5.1 shows the scheme of the adaptive machine learning strategy for catalyst design.
5.3.2 Kinetic analysis

Previously, the mechanism of OER on metals oxides has been extensively studied by the computational method of DFT.[156] Typically, the most favorable reaction pathway in alkaline conditions involves four concerted proton-coupled electron transfer steps: (1) the deprotonation of a water molecule to OH*, (2) the deprotonation of OH* to O*, (3) the production of OOH* by the coupling between O* and a water molecule with a proton-electron transfer, and (4) oxygen evolution by the deprotonation of OOH*. At a given catalyst surface, the free energy diagram of OER involving those reaction intermediates can be constructed using DFT-calculated adsorption energies with tabulated zero-point energy and entropy contributions.[122] Using the computational hydrogen electrode model, the theoretical limiting potential is determined as the minimal operating potential that is required to drive all the elementary steps exergonic. By using the universal scaling relationship[122] established between the free formation energies of OH* and OOH* species ($\Delta G_{\text{OOH}*} = \Delta G_{\text{OH}*} + 3.2$ eV), a simplified OER kinetic model can be established using O* and OH* free formation energies. Fig. 5.2(a) shows the heat map of OER overpotentials ($\eta$) plotted against two activity descriptors, i.e., the Gibbs free energy of OH* and O*−OH* (taken for convenience). To validate the theoretical approach with experimental measurements, we compare the theoretical overpotentials of known ABO$_3$-type perovskites to the experimental activity metrics as shown in Fig. 5.2(b). The activity metric is adapted from the literature[79], which compiles the experimentally measured OER activity of perovskites from different sources normalized with respect to LaCoO$_3$. For the samples with the same composition, the mean is taken. A larger metric value indicates a higher OER activity. It is noteworthy that the DFT+U method is used to describe strongly correlated d-electrons in our study, see Computational Details in supplementary materials for detailed description. However, the DFT+U method predicts a rather large theoretical overpotential for LaCoO$_3$ ($\sim 1$ V), while the traditional
5.3. Results and discussion

DFT-GGA gives a comparable overpotential to experimental measurement (~0.5 V) at the 50 mA/cm² current density.[109, 196] We corrected such a systematic error by applying constant shifts (O*: −0.81 eV and OH*: −0.26 eV) to the DFT+U calculated free energies. A linear trend is observed in Fig 5.2(b), suggesting that the theoretical overpotential (η) is a reasonable indicator toward the OER activity of ABO$_3$-type perovskites.

![Figure 5.2: Theoretical overpotential model and benchmark analysis. (a) Heat map of theoretical overpotential (V) of oxygen evolution using the OH* and O*- OH* free formation energies as descriptors. The known perovskites are overlaid on the volcano. (b) Experimental OER activity metrics vs overpotential (V) for a series of perovskites in alkaline media.](image)

5.3.3 Feature engineering

The material representation is the key ingredient to the development of machine learning models, which requires the in-depth domain knowledge about the material's physical properties. It characterizes the underlying physics of a unique structure with high explainability for the desired properties. Typically, inorganic crystals are mined through three different ways in machine learning, i.e., electronic structure,[111, 112, 121] structural coordination,[87] and atomic properties.[39] In this work, we use a heterogeneous set of material features with
The descriptors include the compounds’ compositional properties and DFT-calculated bulk electronic structure. The 18-dimensional compositional representation is generated by mining the informatics of atomic properties of a constituent metal element that are accessible in the Pymatgen library. More specifically, the compositional features include 4-dimensional atomic properties (i.e., atomic oxidation state and atomic electronegativity) for the cation sites (A and B), 8-dimensional generalized mean of the atomic properties (i.e., atomic number, common oxidation state, Mendeleev number, Pauling electronegativity, thermal conductivity, average ionic radius, thermal conductivity, and outer atomic orbital energy) over the entire structure using the literature method,[39] the tolerance factor and octahedral factor, and 4-dimensional descriptors associating with ionic ratios and electronegativity differences between the A/B cations and O atom using the methods from literature[148]. Besides, 46-dimensional electronic structure descriptors are also included. The computationally inexpensive bulk density of states calculations allow us to quickly assess the material’s orbital-resolved electronic structure of active B-sites and O atom. To transform the semi-structured density of states into a model compatible numeric vector, we calculate the $0^{th}$ - $4^{th}$ moment characteristics (i.e., filling, center, width, skewness, and kurtosis) of the orbital-oriented projected density of states for various atoms in a perovskite bulk structure. More specifically, we calculate the 5-dimensional moment descriptors for each $s$-orbital and 5 individual $d$-orbitals (e.g., $d_{z^2}$, $d_{x^2-y^2}$, $d_{zx}$, $d_{zy}$, $d_{xy}$) of a B-site cation. For the oxygen atom, the moment descriptors are calculated for each of the 3 $p$-orbitals (i.e., $p_x$, $p_y$, and $p_z$). To account for the electronic structure of the $e_g$ molecular orbital, electron occupancy of the $e_g$ orbital is introduced by taking the ratio of the overall filling of the B metal $d_{z^2}$, $d_{x^2-y^2}$ orbitals with respect to the orbitals of $d_{xy}$, $d_{xz}$, $d_{yz}$. All the feature functions are provided in the supporting material and Github. Therefore, a total 64-dimensional inexpensive material descriptors were generated to uniquely represent a perovskite for capturing the reactivity of B-site cations.
5.3. Results and discussion

5.3.4 Model training

In this work, we chose the Gaussian process regression model by leveraging its inherent ability for uncertainty measurement. In definition, Gaussian process is a nonparametric model that defines a probability distribution over a collection of stochastic random functions \((\mu(x_1), ..., \mu(x_n))\) according to the training data \((x_1, ..., x_n)^T\). Generally, the Gaussian process distribution is widely used as a prior function within the context of Bayesian regression. By optimizing the model parameters using the maximum likelihood approach, Gaussian process model can make predictions on unknown feature spaces using the output values of corresponding random functions. Empirically, the values of the random functions are generated in a sampling fashion using a multivariate Gaussian distribution function, which is specified by a mean function and covariance matrix as shown below,

\[
\mu(x_1), ..., \mu(x_n) \sim N(0, K(x_1, ..., x_n)),
\]

\[
K(x_1, ..., x_n) = \alpha^2 \exp\left(-\frac{(x_i - x_j)^2}{2l^2}\right) + \delta_{ij}\sigma^2_{\text{noise}},
\]

where \(\delta_{ij}\) the Kronecker delta for all the possible pairs \((x_i, x_j)\) of data set. A hybrid kernel of the radial-basis function (RBF) and white noise is used in the covariance term, \(K(x_1, ..., x_n)\). As the key component, the kernel function (covariance function) controls the smoothness and the degree of shrinkage toward the mean of the realization of a Gaussian process.[54] In this case, a mixed kernel function, specified by the appropriate parameters such as the scale factor \((\alpha)\), length scale \((l)\), and noise level \((\sigma^2_{\text{noise}})\), would allow an optimal predictive performance of machine learning models.

The prediction performance of the Gaussian process regression model is assessed through
a repeated cross-validation process (bootstrap sampling), where 75% samples are randomly selected from the dataset for training and the model with optimized parameters is applied to predict the rest 25% sample at each iteration. To improve the model stability, all the input features are normalized with 0 mean and the standard deviation of 1 prior to training. In high-dimensional spaces, machine learning models tend to overfit on small datasets, giving a biased prediction for unknown systems. Therefore, a backward feature elimination process is performed by selecting an optimal feature set to reduce the model variance, thus improving the prediction accuracy. Both the averaged training/test RMSE errors are provided in Fig. 5.4(e)(f), and the inserts show the histogram of the training/test error distribution. The uncertainty of model prediction (standard deviation) is labeled as the black bar in Fig. 5.4. The normally distributed errors suggest that the model is validated to explain the underlying dataset corresponding to the basic statistics. The parity plots in Fig. 5.4(e)(f) show that the Gaussian process model optimized with appropriate feature sets can capture the adsorption energies with <0.5 eV errors for both O* and OH*.

5.3.5 Adaptive learning

By learning the structure-reactivity relationships of initially computed 250 perovskites, the machine learning models are applied to explore ~4000 AA’B₂O₆ double perovskites using easily accessible material features, see Fig. 5.3 for specific information of A/B-site elements for the data sets. However, the prediction uncertainty of current models at the initial stage are high because the models are fitted to a small training dataset and have difficulty capturing a large number of unknown samples at the high-dimensional feature space. Therefore, an adaptive learning scheme is designed to systematically adjust the model parameters by explicitly computing candidates with interesting properties albeit large uncertainty. At the initial stage of the adaptive learning process, we approximate the OER overpotentials (η) of ~4000
unknown double perovskites by using the predicted free formation energies ($G_{O^*}$ and $G_{OH^*}$) and kinetic models. As a secondary benefit, the Gaussian process model also estimates the standard deviations ($O^*$ and $OH^*$) for the predicted energies. In this work, we use the averaged value of $O^*$ and $OH^*$ for describing the overall uncertainty ($\sigma_{O^*,OH^*} = (\sigma_{O^*} + \sigma_{OH^*})/2$) of model prediction. During the adaptive learning, we identify the candidate with large prediction uncertainty ($\sigma_{O^*,OH^*} > 1.0$ eV) and small overpotential ($< 0.5$ V) for DFT validation. It is noteworthy that the strategy of picking the next set of candidates for model validation/refinement is the key to adaptive learning. Our selection strategy provides a reasonable trade-off between exploration and exploitation. In the next step, DFT calculations were performed on the selected candidates. The free formation energies and feature descriptors of validated systems are merged into the dataset for model refinement, which is followed by a backward feature elimination procedure for identifying the most appropriate feature sets. In the end, the machine learning model with updated training dataset is applied again to predict the overpotentials for the $\sim$4000 candidates, completing the cycle.

At each iteration, we calculate the ratio of reliable candidates by taking the proportion of the number of reliable samples with respect to the total number of candidates ($\sim$4000). Here, the reliable samples are identified as the samples with prediction uncertainty $\sigma_{O^*,OH^*} < 0.75$ eV. The threshold is taken as the upper bound of model deviations according to our observations. We repeated this adaptive loop until we obtain a ratio of reliable candidates $> 0.8$, so that we gain a much stronger belief of the model reliability. In Fig. 5.4(a)(b)(c), the contour map of the model uncertainties ($\sigma_{O^*,OH^*}$) of $\sim$4000 candidates is plotted against the first two principal components of the original 64-dimensional features. The reduced 2-dimensional coordinates of the 4000 candidates represent the entire materials space in this study. We observe that the reliable region, denoted by the blue color, is expanded gradually to cover the entire 2-dimensional, compressed feature space of $AA'B_2O_6$-type perovskites as
more training data are adopted. Fig. 5.4(d) shows the statistical summary of the model performance in adaptive learning. According to Fig. 5.4(d), the number of promising candidates has some degree of fluctuations versus iterations, which suggests that the original model assumption changes with adaptive learning. The reliable ratio is increased by 40% from the original randomly picked 250 samples to the final iteration, indicating that our model has improved iteratively by learning the structure-reactivity relationships of the candidates and the adaptive learning with the guidance of uncertainty is an effective strategy to improve the machine learning model performance with a small computational effort.
5.3. Results and discussion

Figure 5.4: Adaptive machine learning trends analysis. The uncertainty contour map of model explored candidates versus the PCA transformed 2-dimensional feature space for various training data size of (a) 244, (b) 298 and (c) 358. (d) Summarized statistics of model performance versus training data size of O* and OH* in adaptive learning. The dark blue denotes the ratio of reliable candidates and the grey bar denotes the model identified candidates with promising OER overpotential (< 0.5 V). Note that the training data size at each iteration is the total number of investigated perovskite structures. Parity plots for DFT-calculated vs Gaussian process regression model prediction of adsorption free energies on perovskite surfaces for (e) O* and (f) OH*. The black vertical bars denote the standard deviation of model prediction by Gaussian processes.

5.3.6 Model screening of perovskite electrocatalysts

Ultimately, we chose the optimized machine learning model of the adaptive learning for screening novel perovskite electrocatalysts toward OER. Our model significantly narrows down the chemical space of the AA’B\textsubscript{2}O\textsubscript{6}-type double perovskites (~4000) and identifies a small pool of candidates with near-optimal OER activity and stable cubic structure (tolerance factor > 0.90). To gain compositional insights into the OER activity, we perform an exploratory analysis by comparing the overpotential distribution of the candidates with
different B-site elements. The box plot in Fig. 5.5(a) lists 13 of the candidate categories with the topmost averaged overpotentials (V). Each box summaries a distribution over the overpotentials (V) for a series of perovskites with 5 statistical parameters (e.g., minimum, first quartile, median, third quartile, and maximum). We show that the B-site element type plays a dominant role toward the OER activity and perovskites containing B-site elements of Co, Ru, Ir and Ti have a clustering of candidates with lower overpotentials than other series. Furthermore, the A-site element type also has a secondary effect for the overpotentials from the variations of the quartiles and boundaries of the boxes. Among the promising double perovskites, we are particularly interested in the configurations containing Co and Ti because these elements are earth abundant and inexpensive in comparison with the precious metals such as Ru and Ir. Therefore, special attention is given to Co and Ti for further investigation. The heat map in Fig. 5.5(b) displays the model estimated OER overpotentials along with the tolerance factors for a series of unknown oxide configurations, where the B-site elements are constrained to Co or Ti. Note that the tolerance factor[102] is an indicator of the degree of distortion from an idealized cubic structure and typically a stabilized cubic structure would have a relatively large tolerance factor > 0.9.

To validate our model predictions with respect to experiments, we highlight the model-predicted OER activities for a series of configurations, where the same A/B-site element pairs and ratios have been investigated.[60, 79, 177] For instance, our model shows agreement with the experimental measurements in the literature toward high-performance catalysts by identifying a subset of well-known perovskites (e.g., LaSrCo$_2$O$_6$, LaCaCo$_2$O$_6$, LaCaFe$_2$O$_6$) with a relatively low overpotentials with respect to LaCoO$_3$. Also, some low-performance perovskites with relatively large overpotentials compared to LaCoO$_3$ are also identified, including LaCrO$_3$, LaVO$_3$, LaMnO$_3$, SrVO$_3$, SrTiO$_3$, LaCaMn$_2$O$_6$, LaCaCr$_2$O$_6$ and LaSrCr$_2$O$_6$. The screening result illustrates that the data-driven approach significantly
facilitates the computational screening capability and narrows down the materials space with reasonably low computation effort (≈360 DFT calculated perovskite structures). For guiding the future experiment design, our adaptive learning system recommends the following perovskites: K\(\text{RbCo}_2\text{O}_6\), Ba\(\text{SrCo}_2\text{O}_6\), K\(\text{BaCo}_2\text{O}_6\), K\(\text{CaCo}_2\text{O}_6\), Ba\(\text{PbTi}_2\text{O}_6\), Ba\(\text{RbTi}_2\text{O}_6\), Ba\(\text{SnTi}_2\text{O}_6\), Ba\(\text{TlTi}_2\text{O}_6\), La\(\text{TlTi}_2\text{O}_6\), RbEu\(\text{Ti}_2\text{O}_6\). Furthermore, we perform the DFT calculations to a subset of candidate perovskites. Fig. 5.5(d) suggests that the machine learning models have a reasonable prediction accuracy (≈0.6 eV) compared to the DFT calculations.

Figure 5.5: High-throughput screening of double perovskites enabled by learning model. (a) Box and smarm plots display of the overpotential distributions for the candidates dataset with different B-site metals. The boxes are ordered by the median values for the groups and the horizontal bar denotes the low/high boundary of the spread. (b) Heat map visualization of the OER activity of double perovskites as a function of A-site/B-site cations in terms of the OER overpotentials and cubic phase probability. The red/blue color bar denotes the overpotentials and the purple bar represents the Tolerance factor. (c) Configuration of AA\(\text{'B}_2\text{O}_6\) cubic double perovskite. (d) Parity plots for DFT-calculated vs Gaussian process model prediction of descriptor adsorption free energies on candidate perovskite structures.
5.3.7 Molecular orbital insights

To understand physical factors governing the OER activity, we performed univariate analysis on the candidates with the high/low OER activity. For each material feature, we quantify the dissimilarity of the probability density functions (pdf) with respect to both good and bad candidate datasets using the Kullback-Leibler divergence.\[103\] The Kullback-Leibler (KL) divergence, a quantity that has been developed within the context of information theory, is an asymmetric measurement of the amount of information loss if an empirical probability distribution is approximated by a model distribution. It is also widely used as the objective function in the machine learning models such as the decision tree algorithm.\[73\] A large KL divergence value indicates that the corresponding descriptor is informative and plays a critical role in distinguishing the samples from the two categorical classes. Fig. 5.6(a)(b) show that the pdf distributions of the compositional descriptors, i.e., A-site electronegativity and the tolerance factor, are strongly overlapped for the high/low OER activity samples, resulting in small KL divergence values. In contrast, the electronic structure descriptors in Fig. 5.6(c)(d) are more informative due to the significant mismatch of the probability distributions. In Fig. 5.6(e), we highlight all the topmost informative descriptors (KL divergence > 0.4) and a general conclusion can be drawn that the metal B-site electronic structure descriptors, more specifically the eg orbital ($dz^2$, $dx^2 - y^2$) properties, are closely related to the perovskite OER activity. Furthermore, among all the impactful electronic structure descriptors, the eg electron occupancy, $dz^2$ filling and $dx^2 - y^2$ filling play a dominant role than others with the highest KL divergence values. Importantly, the univariate analysis illustrates that the bulk electronic structure descriptor, as a physically intuitive feature, has a huge advantage to shed light on the underlying physical factors that govern the OER activity from the molecular orbital point of view. Our observations are in agreement with the experimental study[177] that the measured metal B-site eg orbital electron occupancy is correlated with the OER
5.3. Results and discussion

activity. This is because the $d_z^2$ orbital, one component of eg orbitals, is directly overlapping with the p-orbitals of oxygen in reaction intermediates at atop sites.

Figure 5.6: Molecular orbital insights toward OER. The probability distribution plots of high/low OER activity perovskite candidates using the following descriptors of (a) A-site electronegativity, (b) tolerance factor, (c) B-ion $d_x^2 - y^2$ orbital center and (d) $d_z^2$ orbital filling. The grey color denotes the low OER activity candidates and green/blue color denotes the high OER activity candidates with compositional descriptor and electronic structure descriptors. (e) Polar distribution plots of the most informatic descriptors (KL entropy index > 0.4).
5.4 Conclusions

In summary, uncertainty-guided machine learning models are employed to facilitate the discovery of AA’B₂O₆-type double perovskites for the oxygen evolution reaction (OER). By learning the correlations between the easily accessible bulk informatics (i.e., compositional and electronic structure) and the adsorption energies, i.e., O* and OH*, the probabilistic Gaussian processes are able to quickly estimate the reaction intermediates’ adsorption energies and the corresponding uncertainties for a rich material space (≈4000 configurations). To tackle the primary challenge of model overfitting due to data deficiency, we develop an adaptive learning strategy to improve the model reliability progressively by validating those promising samples albeit with large uncertainties. We demonstrated that adaptive learning is a cost-effective solution to shrink the model prediction uncertainties across a high-dimensional feature space with minimum computational cost. The model demonstrated the predictability by identifying some of the literature reported perovskites with measured OER activity. We recommend ~10 stable cubic double perovskites with promising OER catalytic activity that have not been investigated previously. Notably, the physically intuitive nature of the bulk electronic structure descriptors also allows us to draw molecular orbital insights towards perovskites OER activity. By interpreting the probability distributions for the samples from the diverse activity classes with Kullback-Leibler (KL) divergence, we conclude that the $e_g$ orbital electronic structure characteristics of the metal B-site are the underlying factors that govern the OER activity, providing physical insights from data-driven models.
References


[16] Alex Aziz, A Rabdel Ruiz-Salvador, Norge C Hernández, Sofia Calero, Said Hamad,


[36] Stefano Curtarolo, Gus L W Hart, Marco Buongiorno Nardelli, Natalio Mingo, Stefano Sanvito, and Ohad Levy. The high-throughput highway to computational materials design, 2013.


[67] Katja Hansen, Grégoire Montavon, Franziska Biegler, Siamac Fazli, Matthias Rupp,


[82] Jürg Hutter, Marcella Iannuzzi, Florian Schiffmann, and Joost VandeVondele. *cp2k:


[105] D D Landis, J S Hummelshøj, S Nestorov, J Greeley, M Dulak, T Bligaard, J K


[120] Xianfeng Ma and Hongliang Xin. Orbitalwise coordination number for predicting


[128] Joseph H Montoya, Linsey C Seitz, Pongkarn Chakthranont, Aleksandra Vojvodic,


[174] Helge S Stein, Dan Guevarra, Paul F Newhouse, Edwin Soedarmadji, and John M


[182] Zachary W Ulissi, Andrew J Medford, Thomas Bligaard, and Jens K Nørskov. To


