Exploration of Current Methods for Modeling Catalytic Materials: Periodic Density Functional Theory and Machine Learning

A THESIS
SUBMITTED TO THE FACULTY OF THE
UNIVERSITY OF MINNESOTA
BY

Brianna A. Collins

IN PARTIAL FULFILLMENT OF THE REQUIREMENTS
FOR THE DEGREE OF
DOCTOR OF PHILOSOPHY

Jason D. Goodpaster, Adviser

December 2022

Acknowledgements

I have always depended on the kindness of strangers. Over time, with great fortune, these strangers have become family, friends, and loved ones. From the family that took me in when I was temporarily, not so temporarily couch surfing (a.k.a. homeless) to my mom whom I have asked to call me to make sure I am awake for many early morning classes and meetings, there have been many helping hands along my way, and I'd like to thank them all.

Dedication

I dedicate this to my partner, Michael, who also deserves a degree as he has listened to me talk about my research for well over the required number of hours.

1 Abstract

Traditional energy technologies are unsustainable in the modern world. Thus, the development of more advanced technologies is of great significance. From a chemistry perspective, an ideal area for development is within the realm of catalysis. Using catalysts leads to faster and more energy-efficient chemical reactions. Finding an optimal catalyst is of great interest in the communities of chemists and material scientists. To that end, a fundamental understanding of the properties and mechanisms of action of the current most efficient and environmentally sustainable catalysts must be obtained to enable the design of next-generation catalysts. Computational models are very efficient ways of obtaining such information and have become an invaluable component of this work. This dissertation uses computational chemistry methods, namely Density Functional Theory (DFT), to study complex catalytic systems. Data science techniques are also utilized, including machine learning (ML) methods. Many diverse systems are surveyed here, including metal–organic frameworks (MOFs) and surfaces. Section 7 focuses on the computational modeling of magnetic MOFs. Sections 8 and 9 focus on elucidating mechanisms concerning the nitrogen reduction reaction (NRR) and the methane oxidation reaction (MOR) on surfaces, respectively. Section 10 gives a brief introduction on the use of machine learning (ML) techniques in MOFs. Finally, Section 11 uses ML methods to predict adsorption energies on surfaces utilizing nominal information.

2 Table of Contents

3 List of Tables

4 List of Figures

5 List of Abbreviations

KS-DFT	Kohn-Sham Density Functional Theory
DFT	Density Functional Theory
MOF	Metal–Organic Framework
OER	Oxygen Evolution Reaction
MOR	Methane Oxidation Reaction
NRR	Nitrogen Reduction Reaction
HER	Hydrogen Evolution Reaction
TMO	Transition Metal Oxide
SOFC	Solid Oxide Fuel Cells
PBE	Perdew Burke-Ernzerhof
VASP	Vienna ab-initio Simulation Package
LDA	Local Density Approximation
NN	Neural Network
NNP	Neural Network Potential
KRR	Kernel Ridge Regression
SVM	Support Vector Machine
GPR	Gaussian Process Regression
DT	Decision Tree
RF	Random Forest

6 Introduction

The challenge of moving toward a greener and more sustainable society will inevitably require the transformation of many facets of culture and economy. From a green chemistry standpoint, we should strive to make the most efficient use of natural resources and promote substituting fossil fuels with renewable alternatives. Achieving these goals will require the design of novel and efficient chemistries that can be produced sustainably. I aim to utilize modern computational chemistry methods and advances in data science to explore a more sustainable chemical space. To design these next-generation catalysts, we must understand the properties and mechanisms surrounding our current options. Computational models are the most efficient and least resource-heavy ways of obtaining this information and have now become an invaluable component in discovery. Here we choose to utilize methods based in quantum mechanics. The most utilized method to query materials using quantum mechanics is Density Functional Theory (DFT).

Almost all properties of materials can be ascertained by suitable computational tools that solve for quantum mechanical insights. Therefore, we will use the mindset put forward by Dirac in 1929, that progress depends on the development of sufficiently accurate, but tractable, approximate techniques.[1,2] Therefore, the development of DFT and the expression of the tractability and accuracy of the local density approximation (LDA) and various other functionals marked an important milestone in chemistry. LDA, developed and applied by Slater and his co-workers,[3] was progressed before the DFT of Hohenberg and Kohn.[4] However, the usefulness of LDA remained to be seen until the late 1970's when several simulations demonstrating the practicality and accuracy of the approach in assessing the properties of solids appeared.[5,6,7] Since then, the advent of numerous functionals with ever increasing accuracy and practicality, have solidified the use of DFT methods for solids and extended systems. Several excellent reviews on DFT methods exist including those by Emzerhof, Perdew and Burke.[8] The following will consist of a brief overview of DFT.

Taking for example the non-spin-polarized case, DFT states that the total energy, E, is given exactly as a functional of the ground state electronic density, ρ.

$$E = E\left[\rho\right]$$

In this way the true ground state density minimizes the energy, and the other ground state properties of the system are functions of this density. However, there is no theorem to dictate what form this function of energy takes. Therefore, the utility of DFT depends on sufficiently accurate approximations to the form of this function.[5]

One such approximation is Kohn-Sham Theory.[9] Kohn and Sham wrote the electron density as a sum of the single particle densities and then used the variational principle to determine the ground state energy. To do this, Kohn and Sham showed that the correct density is given by the self-consistent solution of a set of single particles Schrödinger-like equations, known as the Kohn-Sham equations, with a density dependent potential (in atomic units):

$$\left(-\frac{1}{2}\nabla^2 + v_s(\mathrm{r})\right)\phi_i(\mathrm{r}) = \varepsilon_i\phi_i(\mathrm{r})$$

Here, ε_i is the orbital energy of the corresponding Kohn–Sham orbital, ϕ_i, $v_s(\mathrm{r})$ is the Kohn-Sham potential that will be discussed below. The density for an N-particle system can be written as:

$$\rho(\mathrm{r}) = \sum_i^N \left|\phi_i(\mathrm{r})\right|^2$$

This allows us to separate the total energy into the following equation:

$$E[\rho] = T_s[\rho] + \int v_{ext}(\mathrm{r})\rho(\mathrm{r})\mathrm{dr} + E_H[\rho] + E_{xc}[\rho]$$

Where T_s is defined as:

$$T_s[\rho] = \sum_i^N \int \phi_i(\mathrm{r})\left(-\frac{1}{2}\nabla^2\right)\phi_i^*(\mathrm{r})\mathrm{dr}$$

and E_H is the Hartree energy (Coulomb energy) and it can be defined as:

$$E_H = \frac{1}{2}\int\int \frac{\rho(\mathrm{r})\rho(\mathrm{r'})}{\left|\mathrm{r}-\mathrm{r'}\right|}\mathrm{drdr'}.$$

Here V_{ext} is the external potential acting on the system. For a molecular system this potential is the electron-nuclei interaction. The Kohn–Sham equations are found by varying the total energy with respect to a set of orbitals, subject to the constraint that the orbitals yield the Kohn–Sham potential as:

$$v_s(\mathrm{r}) = v_{ext}(\mathrm{r}) + \int \frac{\rho(\mathrm{r'})}{\left|\mathrm{r} - \mathrm{r'}\right|}\mathrm{dr'} + \frac{\delta E_{xc}[\rho]}{\delta\rho(\mathrm{r})}$$

Where $v_{xc}(\mathrm{r}) \equiv \frac{\delta E_{xc}[\rho]}{\delta\rho(\mathrm{r})}$ is the exchange-correlation potential. The energy of the system is thus described as:

$$E = \sum_{i}^{N} \varepsilon_i - E_H[\rho] + E_{xc}[\rho] - \int \frac{\delta E_{xc}[\rho]}{\delta\rho(\mathrm{r})}\rho(\mathrm{r})\mathrm{dr}\ .$$

The term E_{xc} is the exchange-correlation energy, which contains the corrections to the kinetic energy due to interacting electrons, the corrections to the electron-electron Coulomb energy from self-interaction, the electron-electron exchange energy, and the electron-electron correlation energy. The simplest approximation to E_{xc} is in LDA, where the functional has been derived from a homogeneous electron gas and is only dependent on the local value of the electron density.[3] In the generalized gradient approximations, an improvement over LDA is made by considering the gradient of the electron density. There are various other types of functionals, however, each functional offers different advantages and disadvantages depending on the nature of the system studied. Thus, the groundwork for KS-DFT methods has been laid out, however, its real utility comes from its full implementation. Therefore, we must discuss basis sets.

Planewave basis sets form a complete set and are a relatively simple basis. In principle the completeness means that relative accuracy can be obtained by increasing the number of planewaves in the basis set, and more importantly that the convergences of a calculation can be monitored by varying the planewave cutoff.[6] Additionally, wavefunctions expanded in planewaves can be transformed efficiently from reciprocal space (coefficients of the planewave expansion) to real space (values on a real space grid) using fast Fourier transforms meaning that many operators can be made diagonal. In particular, the kinetic energy and momentum operators are diagonal in reciprocal space, and the operation of local potentials is diagonal in real space.[1] Therefore, we have chosen to use a planewave basis set because it is complete and computationally efficient.

Thus, instead of having to solve a many-body Schrödinger equation, using KS-DFT in a planewave basis set we have the far easier problem of determining the solution to a series of single particle equations, along with a self-consistency requirement. It is this that makes KS-DFT uniquely poised to provide quantum mechanical level accuracy for a fraction of the cost of traditional wave function methods.[1]

In this dissertation, the study of the chemical properties of complex transition metal catalysis were performed utilizing KS-DFT computational chemistry methods. Section 7 discusses on the computational modeling utilizing KS-DFT of a magnetic MOF in collaboration with the Long Group at UC Berkley. Section 8 focuses on the elucidation of mechanisms concerning the NRR on pure metallic surfaces in collaboration with the Singh group at the University of Illinois at Chicago. Section 9 is concerned with the mechanism of the MOR on metal oxide surfaces, again in collaboration with the Singh group at the University of Illinois at Chicago. Section 10 provides a brief introduction to machine learning within the context of metal-organic frameworks. Finally, Section 11 focuses on the use of machine learning methods to predict adsorption energies on bimetallic surfaces utilizing nominal categorical information.

7 Magnetic ordering through itinerant ferromagnetism in a metal–organic framework

This chapter describes the outcome of a collaborative research project carried out by the authors of the following:

Park, J. G.; Collins, B. A.; Darago, L. E.; Runčevski, T.; Ziebel, M. E.; Aubrey, M. L.; Jiang, H. Z. H.; Velasquez, E.; Green, M. A.; Goodpaster, J. D.; Long, J. R. Magnetic Ordering through Itinerant Ferromagnetism in a Metal–organic Framework. *Nature Chemistry* 2021, *13* (6), 594–598.

Reported is work conducted directly by and in consultation with the author

Reprinted with permission

7.1 Introduction

The development of multifunctional magnets with optimized magnetic properties in concert with other physical properties, such as porosity and high magnetoresistance, remains a daunting challenge, but with the promise to enable new applications in green technologies[10] and next-generation data processing and storage.[11,12] For the realization of commercial applications, it is crucial for a magnet to exhibit a high magnetic ordering temperature beyond potentially elevated operating temperatures. The vast majority of solid-state permanent magnets with high ordering temperatures are based on itinerant magnetism.[13] In particular, a specific form of itinerant magnetism, known as a double-exchange mechanism, was discovered by Zener in 1951 to explain the coexistence of metallic conductivity and high-temperature ferromagnetism in perovskite mixed-valence manganites, which contain Mn(III) and Mn(IV) ions connected by bridging O_2^- ions.[14] Here an itinerant electron is delocalized between the e_g orbitals of neighboring Mn ions and consequently promotes a parallel alignment of spins for the localized electrons in t_{2g} orbitals, in accordance with Hund's rules. Many such double-exchange solid-state materials, which include cobaltites and Heusler alloys, have since been discovered and investigated for applications in spintronics devices.[11,15,16]

As an alternative to traditional solid-state materials, metal–organic frameworks, a subclass of coordination solids, are composed of inorganic building units connected by polytopic organic linkers. Compared with solid-state materials, metal–organic frameworks can offer tremendous synthetic versatility to fine-tune their chemical and physical properties. For instance, organic linkers with predictable binding modes can be utilized to yield frameworks with unique crystal structures and physical properties, which are easily modified through methods such as the substitution of electron donating and/or withdrawing substituents on the ligand, post synthetic redox chemistry and metal or linker exchange.[17,18,19] Furthermore, a combination of a long-range magnetic order and porosity in framework materials could lead to the realization of lightweight permanent magnets and magnetic separation media.[20,21,22] Nevertheless, most framework materials are not permanent magnets, owing primarily to the inability of diamagnetic organic linkers to mediate the strong magnetic coupling requisite for long-range order. Indeed, only a few strategies have been developed to synthesize coordination solids with high magnetic ordering temperatures, which include the employment of short diamagnetic inorganic ligands or organic radical ligands, as exemplified by Prussian blue analogues and the amorphous material V(tetracyanoethylene)~2, respectively.[20,21,23,24,25,26,27] Alternatively, a potentially powerful, yet unrealized strategy, involves the utilization of itinerant charge carriers via a double-exchange mechanism. In addition to achieving high ordering temperatures in metal–organic magnets that contain diamagnetic linkers, this approach may further provide a means to introduce metallic electronic conductivity.

Although double exchange is a recognized phenomenon in solid-state materials,[14,15,16] examples in coordination solids with organic ligands are limited to molecular compounds.[28,29] Azolate ligands have strong σ-donating and π-accepting abilities, and when coordinated to octahedral metal ions with diffuse dπ orbitals of favorable energies, strong π–d conjugation between the ligand and metal orbitals may arise.[28,29,30]

Furthermore, compact, symmetrical azolate ligands can support crystal structures that consist of infinite metal–azolate chains with short metal–metal distances and an octahedral coordination environment around the metal ions,[31] which provides efficient pathways for long-range charge transport and magnetic interaction.[32,33,34] Here we report the mixed-valence framework material $Cr(tri)_2(CF_3SO_3)_{0.33}$, which exhibits itinerant ferromagnetism with a Curie temperature (T_C) of 225 K via a double-exchange mechanism.

7.2 Computational Details

KS-DFT calculations were performed using the Vienna ab-initio Simulation Package (VASP) version 5.3.5.[35] The electronic wave-functions were expanded in a plane-wave basis set with an energy cutoff of 400 eV. Electron-ion interactions were described using the projector augmented wave method in the form of pseudopotentials found in the VASP library.[36] All structures were fully relaxed using the ISIF = 3 tag in VASP at a 1×1×1 Monkhorst-pack k-point mesh and a force tolerance of 0.02 eV/Å using the Perdew Burke-Ernzerhof (PBE) functional.[37] The low k-point sampling for geometry optimizations was required due to the large unit cell; however, subsequent self-consistent calculations for the energies and charge densities were performed at a 3×3×3 Monkhorst-pack k-point mesh and the PBE functional with an energy tolerance of 10^{-5}. The charge analysis was performed using the Chargemol DDEC6 scheme.[38,39,40] The resultant structures with magnetic moment and atomic charge color gradients were constructed in

Jmol using the information from the DDEC6 calculations.[41] All additional visualization of the systems was done in Visualization for Electronic and Structural Analysis (VESTA).[42]

7.3 Results and Discussion

We generated the geometry to use for our KS-DFT calculations starting from the crystal structure of $Cr(tri)_2(CF_3SO_3)_{0.33}$. The crystal structure was disordered around the counterions, therefore the initial position of the counterions were estimated and hydrogen atoms were added to fill missing valences. The resultant structure has 424 atoms. The counterions were kept in all calculations. Experimental data suggest disordered triflate anions. We only tested a limited number of triflate configurations. Our results suggested that there was not a large impact on the magnetic state; however, the exact impact on anion configuration on the magnetic state requires further study. The PBE functional was used for the optimization and subsequent calculations because this functional is free of empirical parameters and gives good results for the structure of densely packed solids.[37] Although previous work on magnetic metal-organic frameworks used the HSE functional,[43] this would have been too computationally expensive for a system of this size. Furthermore, we did not include a U parameter correction in our calculations. A U parameter is frequently included to correct over-delocalization seen in KS-DFT calculations.[44] However, such a correction is typically not included for systems with delocalized electrons, such as metals and conductors. Experimentally,

$Cr(tri)_2(CF_3SO_3)_{0.33}$ shows high conductivity at low temperatures; therefore, a U parameter is not required. We did test applying the U parameter, and we found a decrease in the mixed-valent character, with less charge difference between different Cr centers. Additionally, applying a U value increases the magnetic moment of our calculations with high values of U (8.0 eV) approaching 2.66 μB, while omitting a U value gives 2.41 μB, in good agreement with the experimentally observed value of 2.39 μB.

Figure 1 exhibits the convergence of the energy with respect to the plane-wave energy cutoff and irreducible k-points. We find that a plane-wave cutoff of 400 eV is sufficient with less than 0.1 eV energy change compared to 425 eV. Additionally, we find that 3×3×3 k-point sampling to be sufficient. The energetics taken from subsequent self-consistent energy calculations suggest that the ferromagnetic state is lower in energy than the antiferromagnetic state by 0.127 eV/Cr atom. In all calculations, the Fermi smearing technique was used to determine orbital occupancy with a width of 0.1 eV. This leads to a broadening of the DOS compared to the band structure by approximately 0.1 eV.

The pDOS (Figure 3d) and Figure 2 were obtained using data that was parsed using Pymatgen (Python Materials Genomics), which is an open-source Python library for materials analysis, and then plotted using Matplotlib.[45,46] The band structures (Figure 3(a,b)) were obtained using Sumo, a Python toolkit for plotting and analysis of *ab initio* solid-state calculation data.[47]

Figure 2 shows the net atomic charge and the even-tempered magnetic moments as a function of color gradient from blue and white to red for the ferromagnetic (Figure 2 (a,b)) and the antiferromagnetic states (Figure 2 (c,d)), respectively. In the Figure 2(a,b), there is clear evidence that the state captured is ferromagnetic and that the distribution of magnetic moments on the Cr atoms ranges from around two to three unpaired electrons, which is consistent with experimental evidences for mixed-valence low-spin Cr(II/III). The same conclusions can be drawn from the net atomic charges. The Figure 2(c,d) also distinctly displays confirmation that the state captured is antiferromagnetic while the net atomic charges show the expected mixed-valency between Cr(II/III). These findings are consistent with the experimental evidence for the $Cr(tri)_2(CF_3SO_3)_{0.33}$ and support double-exchange phenomenon.

Interestingly, the Figure 2(c) shows that the antiferromagnetic state consists not of alternating spin orientations between nearest-neighbor Cr atoms but of several spin-up and spin-down ‘domains’ within a unit cell. Like the ferromagnetic state, the antiferromagnetic state also exhibits significant hybridization between the Cr(d) and N(p) orbitals.

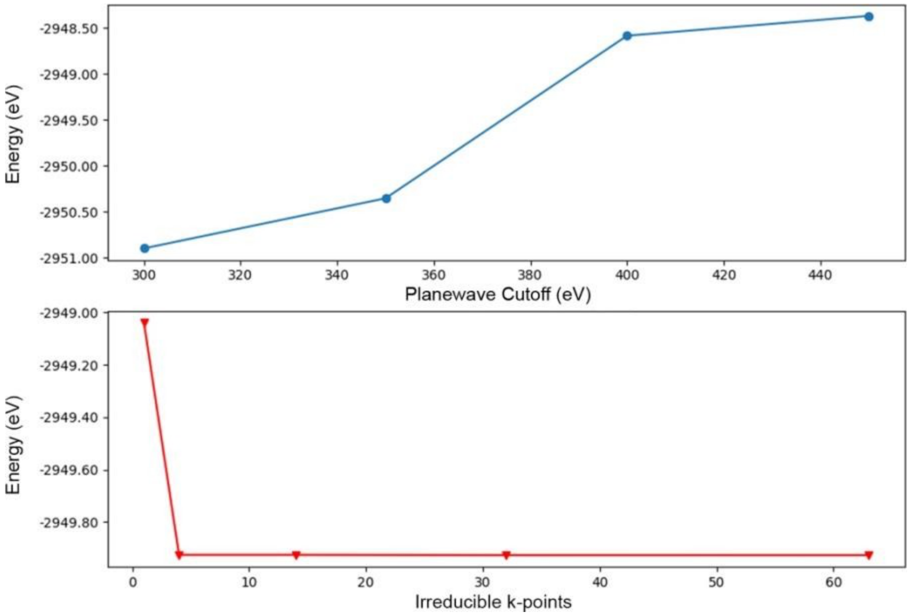


Figure 1 : Energy calculations. The energy of the optimized system is shown with respect to planewave cutoff and number of irreducible k-points. The calculations done varying planewave cutoff include 14 irreducible k-points.

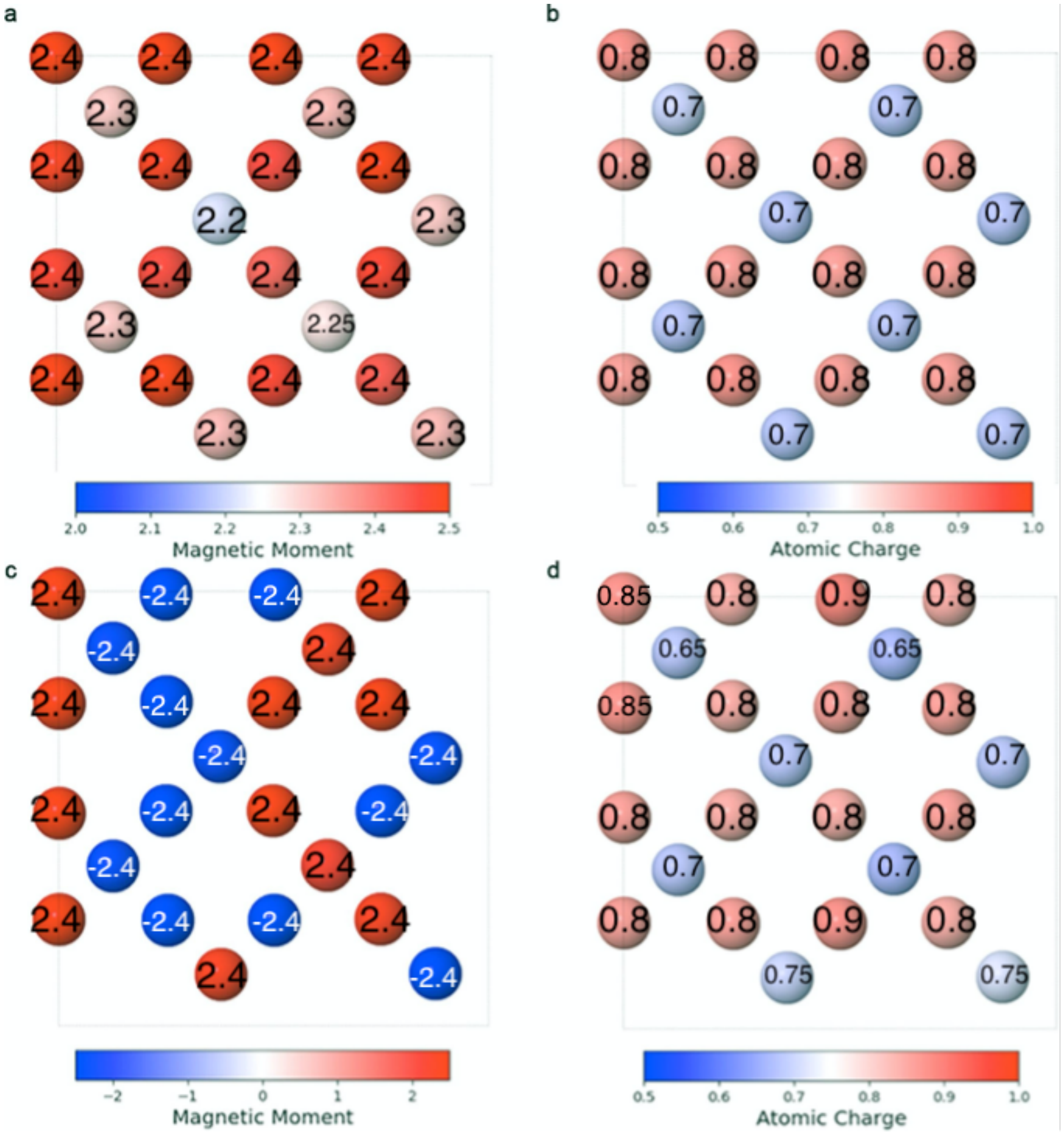


Figure 2 : Charge analysis. a, Magnetic moments (μ mol^{-1}) for the ferromagnetic state. b, Atomic charge for the ferromagnetic state in B atomic units. c, Magnetic moments (μ mol−1) for the antiferromagnetic state. d, Atomic charge for the antiferromagnetic state in atomic units. Atomic B properties shown as a gradient from blue to white to red. Negative magnetic moments indicate spin-down contributions.

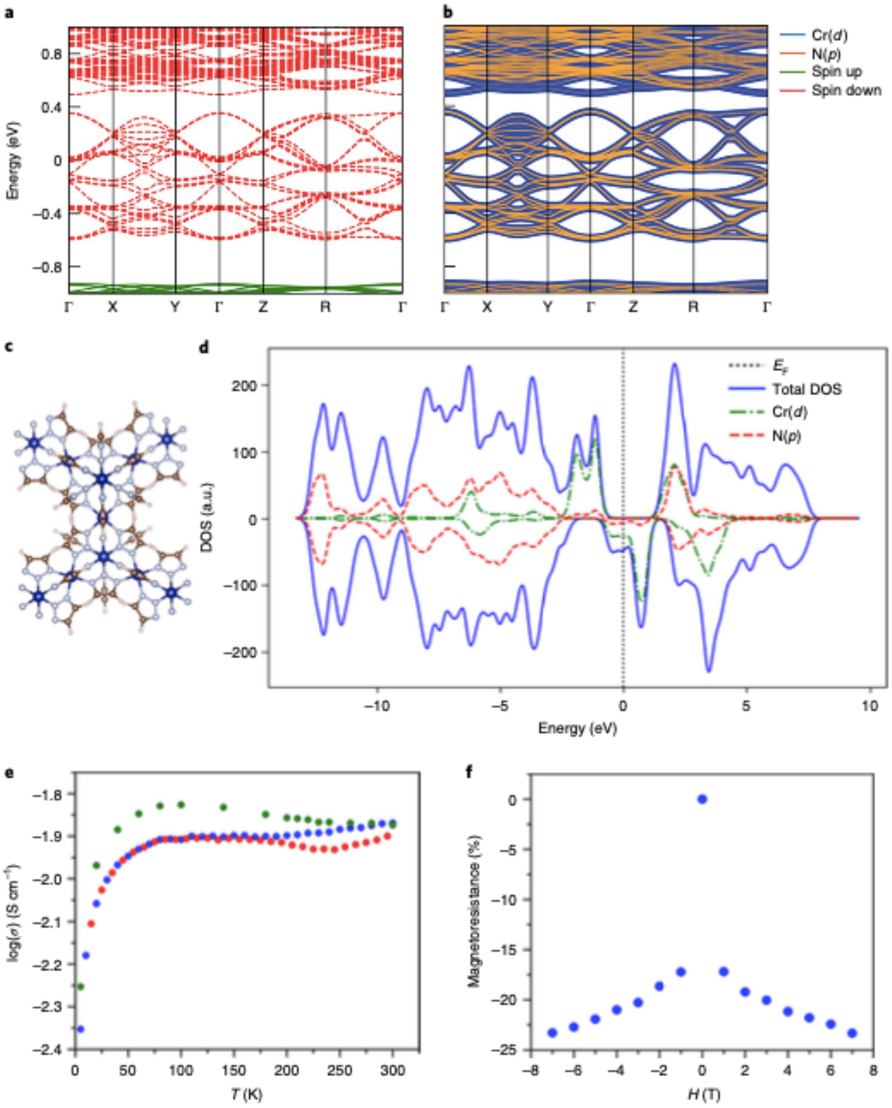


Figure 3 : Electronic structure calculations and electronic conductivity of $Cr(tri)_2(CF_3SO_3)_{0.33}$. a,b, Total (a) and projected (b) band structure for the ferromagnetic state, calculated using the Perdew–Burke–Ernzerhof functional. The zero energy is set to the highest occupied state. The symmetry points correspond to Γ = (0,0,0), X = (0.5,0,0), Y = (0,0.5,0), Z = (0,0,0.5) and R = (0.5,0.5,0.5). The dense spin-down bands near the Fermi energy (E_F) have dominant contributions from the strongly hybridized Cr $d\pi$ and N p orbitals. **c**, The geometry-optimized structure of $Cr(tri)_2(CF_3SO_3)_{0.33}$. Cr, blue; N, light blue; C, brown; O, white; the CF_3SO_3 ion is omitted for clarity. **d**, Partial density of states (DOS) for the ferromagnetic state. Zero energy is set to 33 the highest occupied state, denoted by the dotted vertical line. **e**, Variable-temperature conductivity data of $Cr(tri)_2(CF_3SO_3)_{0.33}$. Blue and red spheres represent measurements taken during cooling and warming, respectively. Green spheres represent data collected while cooling under an

applied magnetic field of 7 T. The increased conductivity under a magnetic field below T_C is consistent with a negative magnetoresistance. **f**, Magnetoresistance data of $Cr(tri)_2(CF_3SO_3)_{0.33}$ collected at 5 K under selected d.c. magnetic fields with a maximum negative magnetoresistance of ~23% at 7 T. a.u., arbitrary units.

The calculations suggest that the ferromagnetic state is lower in energy compared with that of the antiferromagnetic state by 0.127 eV per Cr atom, in agreement with the magnetic measurements. In the ferromagnetic state, the calculated saturation magnetic moment is 2.41 μB, which is in excellent agreement with the experimentally measured value of 2.39 μB. The band structure of $Cr(tri)_2(CF_3SO_3)_{0.33}$ (Figure 3a) near the Fermi energy predominantly consists of spin-down bands with a large density of states and a band dispersion of ~1 eV, which suggests substantial spin polarization. Here, spin polarization is defined as the extent to which the spin is aligned in a particular direction.[48] The projected band structure (Figure 3b) and projected density of states (Figure 3d) show that both Cr(d) orbitals and N(p) orbitals of the triazolate ligands contribute to the valence band maximum and the conduction band minimum. Near the Fermi energy, the frontier orbitals of the bands are dominated by the Cr(dπ) orbitals and N(p) orbitals, with an approximately 32% hybridization between them, which indicates a strong hybridization between the Cr(d) and N(p) orbitals and π–d conjugation. Figure 3d also highlights spin polarization, as illustrated by the substantial peak in the spin-down density at the Fermi energy for both the total and the Cr(d) orbitals. This shows that the Cr(d) orbitals primarily contribute to the spin-polarized state. The localized charge density (Figure 2) shows alternating high and low charge Cr centers. The Cr centers of lower and higher charges exhibit lower and higher magnetic moments, respectively, consistent with the mixed-valence low-spin Cr(II/III) centers. Notably, the charge density and projected density of states both indicate spin polarization. Taken together, these computational results demonstrate a ferromagnetic ground state for $Cr(tri)_2(CF_3SO_3)_{0.33}$, and the high density of spin-polarized bands near the Fermi energy that arises from a strong π–d hybridization suggests the delocalization of electrons in the valence band maximum, is consistent with a double exchange that stems from the mixed-valence metal centers.

7.4 Conclusion

The foregoing results demonstrate the observation of itinerant ferromagnetism in a metal–organic framework, which gives rise to a magnetic ordering temperature of T_C = 225 K and a large negative magnetoresistance. Additionally, this synthetic approach of using π–d conjugation between organic linkers and mixed-valence metal centers with diffuse dπ orbitals provides a general blueprint from which to design materials with improved magnetic and charge-transport properties, given the vast chemical versatility offered by metal–organic frameworks.

8 Competing Effects of pH, Cation Identity, H_2O Saturation, and N_2 Concentration on the Activity and Selectivity of Electrochemical Reduction of N_2 to NH_3 on Electrodeposited Cu at Ambient Conditions

This chapter describes the outcome of a collaborative research project carried out by the authors of the following:

Kani, N. C.; Prajapati, A.; Collins, B. A.; Goodpaster, J. D.; Singh, M. R. Competing Effects of pH, Cation Identity, H_2O Saturation, and N_2 Concentration on the Activity and Selectivity of Electrochemical Reduction of N_2 to NH_3 on Electrodeposited Cu at Ambient Conditions. *ACS Catalysis* **2020**, *10* (24), 14592–14603.

Reported is work conducted directly by and in consultation with the author.

8.1 Introduction

The electrochemical reduction of dinitrogen (N_2) and H_2O to ammonia (NH_3) is of exceptional scientific, societal, and industrial importance.[49] Such a route to produce NH_3 can effectively store and carry hydrogen, reduce the carbon footprint due to the Haber-Bosch process, balance the nitrogen cycle by fixing atmospheric N_2, and provide means to produce on-demand fertilizers using air, H_2O, and sunlight. However, N_2 is a highly stable molecule with a strong N≡N triple bond that makes it extremely difficult to activate at ambient conditions. Applying an electric potential to an electrocatalyst can substantially reduce the activation barrier for N_2 reduction reaction (NRR) for the synthesis of NH_3 at ambient conditions.[50] The outstanding challenge is to minimize the over-reduction of the proton source- H_2O in the hydrogen evolution reaction (HER) while promoting the NRR. Here, we performed a theory-guided study for the search of efficient NRR catalysts and identified strategies to increase activity and selectivity of NRR by optimizing the composition of inner Helmholtz plane by varying electrolyte pH, cation-type, H_2O saturation, and dissolved N_2 concentration.

Although the difference in the equilibrium potential between HER and NRR (shown below) is minimal ~57 mV, the HER is kinetically dominant than NRR for most catalytic systems.

(NRR) $$N_2 + 6H_2O + 6e^- \rightarrow 2NH_3 + 6OH^- \qquad E^0 = 0.057\ \text{V vs. RHE}$$

(HER) $$2H_2O + 2e^- \rightarrow H_2 + 2OH^- \qquad E^0 = 0\ \text{V vs. RHE}$$

This is because the concentration of H_2O is at least three orders of magnitude higher than the solubility limit of N_2 – 1.3 x 10^{-3} mol L^{-1} in H_2O,[51] and the binding energy of H_2O/H on most transition metals are also higher than N_2.[52, 53] The lower solubility and lower binding energy of N_2 in aqueous electrolytes are the primary cause for much lower coverages of NRR intermediates and thereby activity and faradaic efficiency (FE) of NH_3 on planar electrodes (see schematic in Figure 4(A)). Broadly, three different hierarchical approaches can be applied to improve the activity and FE of NRR, namely, design of catalyst, engineering of electrolyte, and optimization of the electrochemical cell.

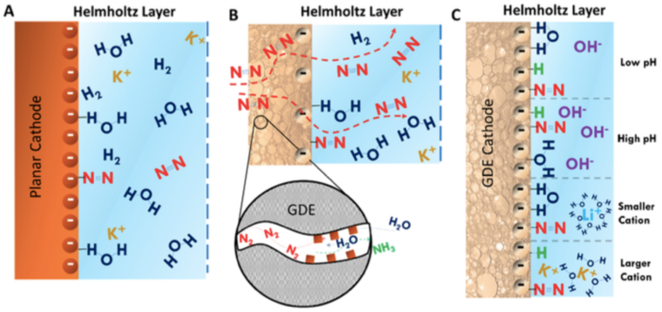


Figure 4: Distribution of H_2O and N_2 near the electrocatalyst in (A) planar electrode and (B) gas-diffusion or porous electrode configurations for aqueous electrochemical cells. The surface coverage of N_2 in the planar electrode configuration is limited by its solubility in H_2O, whereas N_2 coverage in the gas-diffusion electrode can be controlled independently by adjusting the pressure of N_2 at the backside of the electrode. (C) Effect of increasing pH and cation size on H_2O reorganization, *H binding, and stabilization of intermediates. Increasing the pH increases the *H coverage and reorients H_2O to the O-down position on the cathode. Increasing the cation size stabilizes the NRR intermediates and allows direct reduction of H_2O in the solvation shell of larger cations.

The first approach is primarily guided by KS-DFT for the discovery of efficient NRR catalysts that surpass the fundamental limits of existing catalysts for the scalable synthesis of NH_3 at ambient conditions. Extensive KS-DFT studies have been performed to study NRR.[54,55] However, most only focus on the binding energies of the intermediates in a slab/vacuum model[55,56,57,58,59,60] and do not consider the effect of liquid electrolyte, applied potential, and coverages of *H. A recent KS-DFT study reported a kinetic volcano of activation barriers representing a trade-off between weak versus strong N-binding transition metals for optimal NRR.[61] They showed early transition metals of high N-binding favor dissociative mechanism with hydrogenation of *NH as the rate-limiting step; whereas the late transition metals of lower N-binding support associative mechanism with hydrogenation of *N_2 as the rate-limiting step.[13] Similar trends in NRR activity of transition metals have been obtained by calculating the free energy change of elementary steps.[62,63] However, these KS-DFT studies have also reported higher activity of HER on transition metals as compared to NRR, demonstrating an extreme challenge for the identification of efficient NRR catalysts. Besides consideration of the relative activation barriers for HER and NRR, the relative coverages of *H and *N_2 are also crucial in determining the activity and FE of NRR. Therefore, we hypothesize that the most efficient NRR catalysts are the materials with lower HER activity (higher activation barrier) and lower *H coverages, while also possessing reasonable energy barriers for NRR. Here we evaluate NRR activity of the group-11 elements such as Cu, Ag, and Au that have lower HER activity[61] and lower *H coverages (<10%).[64,65] This is accomplished by identifying a new hypothesis-driven descriptor for efficient NRR catalyst. Here presented are the theoretical contributions, including KS-DFT calculations for elementary steps of NRR on Cu, Ag, and Au.

8.2 Computational Details

The estimated energies of the partial mechanism of NRR on the Cu(111), Au(111), and Ag(111) surfaces have been calculated with KS-DFT using the Vienna *Ab initio* Simulation Package (VASP) 5.3.5.[35] Here we perform the calculations simply to determine the activity difference for these metals. Thus, we employ a simplified model that does not consider a fixed potential or explicit solvation. While the absolute energies might have errors, we expect the trend in activity to be accurate. These calculations were performed using the RPBE exchange-correlation functional.[57] Here, the metal surface was modeled using a three-layer (3 x 3) periodic slab, and the successive slabs were separated by at least 20 Å of vacuum. Adsorption was allowed only on one side of the slab. Partial occupancies for each orbital have been populated using the Fermi smearing method with a width of 0.1 eV. Convergence of total energy with respect to planewave cutoff and the k-point set were considered (see Figure 5). Energies reported were sampled at a 4 X 4 X 1 Monkhorst−Pack k-points and a 500 eV planewave cutoff, which, given the information in Figure 5, we found appropriate.

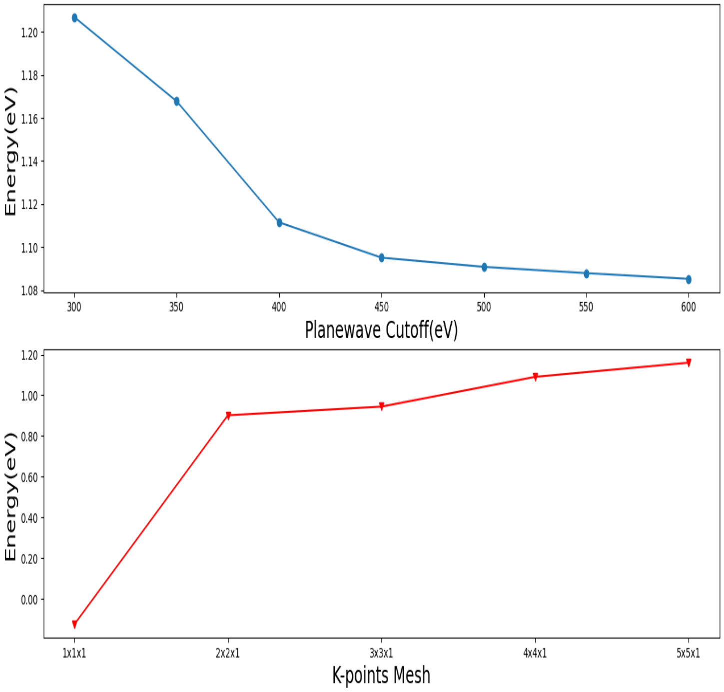


Figure 5 : The energy of the N_2 adsorption reaction on the Cu(111) surface is shown with respect to plane wave cutoff and k-point mesh The calculations done on varying plane wave cutoff include a 4x4x1 k-point mesh. The calculations that vary k-point mesh are at a plane wave cutoff of 500 eV.

8.3 Results and Discussion

$$* + N_{2(g)} \rightarrow * N_2$$

$$* N_2 + \frac{1}{2} H_{2(g)} \rightarrow * N_2H$$

$$* N_2H + \frac{1}{2} H_{2(g)} \rightarrow * NNH_2$$

We postulate that the most efficient NRR catalysts are the materials with lower HER activity (higher overpotential) and lower *H coverages while also possessing reasonable energy barriers for NRR. Catalysts with lower energy barriers for NRR and lower HER activity, but higher *H coverages will also not yield higher NRR activity due to the lack of open binding site for N_2 to bind. Here we evaluate NRR activity of the group-11 elements- Cu, Ag, and Au which have lower HER activty[13] and lower *H coverages (<10%).[64,65,66] The preferred mechanism of NRR over Cu, Ag, and Au is the associative distal pathway, where the hydrogenation of $*N_2$ has the highest barrier and is the rate-limiting step, which is shown in Figure 6. From Figure 6, it is apparent that the energy change for the rate-limiting step increases in the order Cu (111) < Ag (111) < Au (111), which indicates Cu is an effective catalyst for NRR.

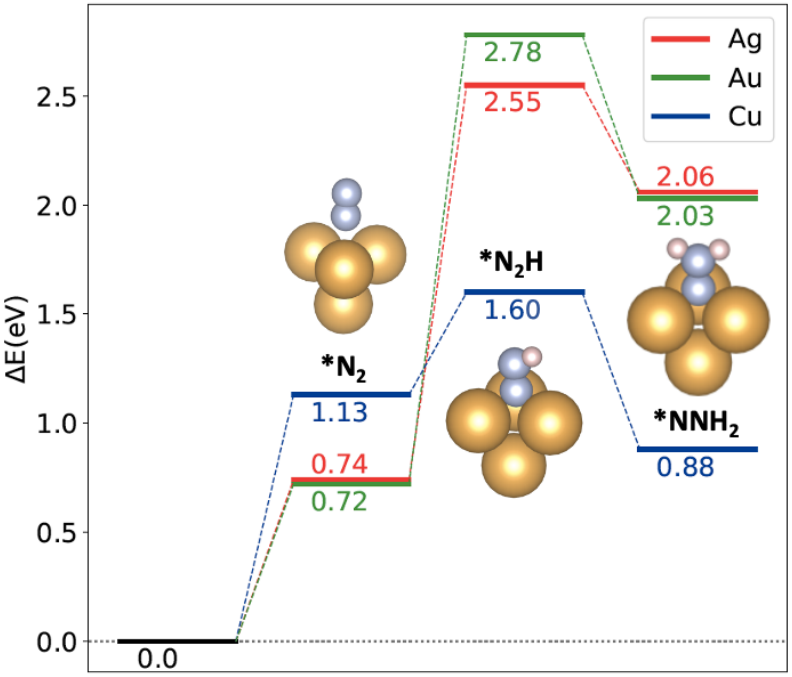


Figure 6 : Reaction energy profile of the partial associative distal mechanism of NRR over Cu(111), Ag(111), and Au(111) at zero-bias.

8.4 Conclusion

In conclusion, we developed a rational approach is developed to design catalysts, electrolytes, and electrochemical cell to obtain higher FE and current density of NH_3. This approach involves a hypothesis-driven descriptor to identify efficient NRR catalysts. The present study provides deeper insights into the effects of the catalysts structure and the reaction environment on the rates of NRR. These insights will generally be applicable to other catalytic materials and will guide the development of efficient systems for electrosynthesis of NH_3.

9 Fundamental insight into electrochemical oxidation of methane towards methanol on transition metal oxides

This chapter describes the outcome of a collaborative research project carried out by the authors of the following:

Prajapati, A.; Collins, B. A.; Goodpaster, J. D.; Singh, M. R. Fundamental Insight into Electrochemical Oxidation of Methane Towards Methanol on Transition Metal Oxides. *Proceedings of the National Academy of Sciences - PNAS* 2021, *118* (8).

Reported is work conducted directly by and in consultation with the author.

9.1 Introduction

Electrochemical oxidation of methane (CH_4) at ambient conditions offers a sustainable route for efficient utilization of abundant natural resources such as shale gas and biogas. However, the lower activity and selectivity of current electrocatalysts pose hurdles for the large-scale deployment of electrochemical technologies for the efficient utilization of CH_4.[67] Currently, the majority (~66%) of CH_4-rich resources are burned to produce electricity or to provide heating for the residential and commercial buildings, which contributes ~1 gigaton of CO_2 emissions annually. CH_4 is also utilized to produce oxygenated chemicals such as CH_3OH using industrial processes like steam reforming followed by gas-phase conversion, or direct thermocatalytic conversion.[68] While the thermocatalytic routes often require high temperature and pressure and suffer from catalyst poisoning[69,70,71] the electrochemical technologies offer environmentally benign and sustainable routes for storing electrical energy by converting CH_4 and H_2O to CH_3OH, and for generating electrical energy using a direct CH_4 fuel cell that primarily generates CO_2.[70,72,73] However, the primary challenge in such electrochemical processes is the first step of CH_4 activation on electrocatalysts, which is difficult at ambient conditions owing to its high C-H bond energy of 439 kJ/mol,[74] high symmetry with tetrahedral molecular geometry, low polarizability of 2.448 $Å^3$,[75] the low solubility of 1.272 mM in the water at STP[76], and competitive oxygen evolution reaction (OER).[77,78] A complete understanding of the mechanism of electrochemical activation followed by oxidation of CH_4 and its competition with OER on transition metal oxides has not been developed yet. Consequently, there is a need for the experimental and theoretical evaluation of the activity and selectivity of methane oxidation reaction (MOR) on various transition metal oxides to elucidate the origins of higher activation energy and competitive kinetics of MOR and OER, and to identify strategies for selective synthesis of CH_3OH under ambient conditions.

There are very few experimental reports available on the kinetics of electrochemical MOR at ambient conditions.[79] The majority of the MOR work focuses on the high-temperature electrocatalysis in galvanic cell configuration, such as solid oxide fuel cells (SOFCs), where the primary objective is to harvest electrical energy by fully oxidizing CH_4 to CO_2.[80,81,82,83,84] The conventional SOFCs operating at a temperature range of 300-700 ^{o}C [83,85] using Ni-based composite anode[86,87,88] have also been studied for partial oxidation of CH_4 to hydrocarbons (e.g., CO, C_3H_4, C_2H_6, and CH_3OH), but their operating efficiencies drop rapidly due to carbon deposition (coking).[72,85,89] The low-temperature electrolytic systems operating at temperatures <120 ^{o}C using Pt, Platinized-Pt, or Pt/Au catalysts have also been reported for partial oxidation of CH_4 to CH_3OH, but the faradaic efficiencies (FE) are too low for a practical purpose.[77,90,91,92] A comprehensive table of the current density, FE, and reaction products for MOR on various electrocatalysts along with electrolyte composition, temperature, and operating potential are provided in Table 1. Insofar, only precious metals like Pt, Pd, Ru, and Au have been explored for MOR under either harsh chemical conditions or at very high overpotentials.[77,93,94,95] Further experimental evaluation of activity, selectivity, and mechanism of MOR over other transition metal catalysts is required to identify efficient MOR catalyst and uncover fundamental limits of CH_4 activation on these catalytic systems. It is also desirable to develop a theoretical understanding of the competitive mechanism of MOR and OER on these transition metal oxides and identify the non-competitive routes for partial oxidation of CH_4.

Table 1 : Table for Aqueous electrochemical systems for methane oxidation

No.	**Catalyst**	**Electrolyte**	**Products (FE)**	**V vs RHE**	**Current density (mA/cm^2)**	**Ref**
1	Pt	CH_3CN		3.87	-	[93]
2	Pt	KCl (pH 11)	CH_3OH, CH_3Cl, CH_2Cl_2, $CHCl_3$, CCl_4 (total FE:18%)	2.149	6.16	[96]
3	Co-Phthalocy-anine on C	1M H_2SO_4+5mM $FeSO_4$	HCHO(20%),HCOOH(5-8%), CH_3OH+ EtOH(3%)	-0.619	30	[90]
4	Hg/Cu	2M KOH	HCHO(36%), CH_3OH(1.2%)	0.8	0.1	[94]
5	Pt, Ru, Pd, Au	0.5M $HClO_4$	CO_2	0.56-0.9	-	

6	Pt	Bis(trifluoromethylsulfonyl)imide-based ionic liquids	CO_2	0.9	0.4	[97]
7	Pt, Pd on C	0.5M H_2SO_4	CO_2	1.2	0.8	[98]
8	Pt	0.1 M $HClO_4$	CO	0.5	30	[99]
9	Pt salts	10mM NaCl+0.5 M H_2SO_4	CH_3OH, CH_3Cl, HCOOH, CO_2	0.89	1.1	[100]
10	MO_x (M=transition metals)	KOH	CH_3OH, CO_2	-	10	[101]
11	NiO-ZrO_2	0.1M Na_2CO_3	CH_3OH, HCHO, CO	2.24	21	[78]
12	Co_3O_4-ZrO_2	0.5M Na_2CO_3	CH_3OH, HCHO, C_2H_4O, C_3H_8O, C_3H_6O (total FE:47%)	2	4.7	[102]

13	TiO_2-RuO_2/PTFE	0.1M Na_2SO_4	CH_3OH, HCHO, HCOOH	2.54	13	[103]
14	Ni-Foam	1M KOH	CH_3OH, C_2H_5OH	1.57	62	[79]
15	FTO	$PdSO_4$ in H_2SO_4	CH_3OSO_3H, CH_3SO_3H	1.29	4.5	[104]

Theoretical investigations of CH_4 adsorption via C-H bond activation followed by electrochemical oxidation are necessary to identify the activity descriptors and fundamental scaling relationships for MOR on transition metal oxides.[105] CH_4 being a stable, non-polar molecule, is known to bind weakly on the transition metals through a dissociative mechanism,[106] whereas all the CH_x ($x = 1, 2$, and 3) intermediates bind more strongly. Psofogiannakis and coworkers[48] showed that CH_4 oxidation to CO is preferable through the dissociative chemisorption of CH_4 on Pt with the sequence $*CH_3 \rightarrow *CH_2 \rightarrow *CH \rightarrow *CHOH \rightarrow *CHO \rightarrow *CO$. A similar study showed, depending on the surface structure, the oxygen assisted dehydrogenation drives the selectivity of the reaction to either CH_xO_y or CO_x products.[107] The presence of oxygen on transition metals in the form of metal oxides has been determined to play a key role in C-H bond activation the oxidation of CH_4.[108,109,110] Under the anodic overpotentials, the dissolved CH_4 dissociates to form *O-CH_3 (CH_3 bonded to metal oxide) and *O-H. Due to H-C-H bond symmetry in CH_4 and CH_3, the energy needed to break (or activate) the C-H bond is equal to the energy required to transfer H from CH_4 to *O (metal oxide). Therefore, the suggested activity descriptor for MOR is E_{OH}-E_O. Whereas the activity descriptor for OER based on the rate-limiting step (*OH → *O) is E_O-E_{OH}.[111,112,113,114] The participation of *O species in both MOR and OER is the primary cause for the competitive kinetics determining the selectivity of CH_4 oxidation on transition metal oxides.

Regardless of the catalyst structure and composition, there has been an agreement in acknowledging that MOR proceeds through a dissociative mechanism. However, in the oxidation regime of the electrocatalytic MOR, the metals exist in the form of their oxides, and one descriptor (either E_H or E_{OH}-E_O) may not be sufficient to identify efficient catalyst as the MOR activity can also depend on the binding energy of the CH_x to the metal oxide surface, and the availability of metal oxide sites to bind with CH_x and consequently to suppress OER.[114,115] Rossmeisl and coworkers[114] elucidate this further by studying the limitations of CH_3OH production on different transition metal oxide (110) surfaces. They show that activation energy for MOR is correlated with the reactivity of an adsorbed oxygen atom from the electrolyte. A catalyst is providing more reactive oxygen results in low energy barriers for OER, which leads to the suppression of MOR. Here we hypothesize that the higher MOR activity requires higher values of E_{OH}-E_O, the higher surface coverage of *O, and a non-competitive binding site of *O. Hence, studying the activity descriptors in combination with the estimation of the binding energy of the CH_x species, the surface coverage of oxygen ad-atoms, and the availability of competitive vs. non-competitive sites for CH_x adsorption on transition metal oxides are expected to provide mechanistic insights to identify active electrocatalysts.

We recognize the experimental challenges associated with C-H bond activation and mass transfer limitations of CH_4 in aqueous electrolytes, as well as theoretical shortcomings due to lack of comprehensive information about the MOR mechanism over transition metal oxides. To address these challenges, we conduct controlled electrochemical experiments to study the activity and selectivity of CH_4 oxidation over 12 transition metal oxides: Sc_2O_3, TiO_2, ZrO_2, Fe_2O_3, Co_3O_4, IrO_2, NiO, PtO_2, Cu_2O_3, ZnO, SnO_2, and PbO_2 in near-neutral and alkaline electrolyte using rotating-disk electrode cell, followed by theoretical evaluation of efficient MOR catalysts using KS-DFT. The objectives of this study are to relate the activity of MOR to the binding energy of CH_4, identify the non-competitive active sites for MOR-active catalysts, and to determine reaction pathways and strategies to direct partial oxidation of CH_4 to CH_3OH.

9.2 Computational Details

Estimated energies of adsorption and the free energy profile on four metal oxide surfaces, IrO_2, PbO_2, TiO_2, and SnO_2 have been calculated with KS-DFT in the Vienna *Ab initio* Simulation Package (VASP) 5.3.5.[35] Calculations were performed using the PBE exchange-correlation functional with Grimme's D3 corrections.[116,37] Here, the metal oxide (110) surfaces were modeled using a four-layer (4 X 4) periodic slab, and the successive slabs were separated by at least 25 Å of vacuum. Adsorption was allowed only on one side of the slab. Partial occupancies for each orbital were populated using the Fermi smearing method with a width of 0.1 eV. All energies reported were sampled at a 4 x 4 x 1 Monkhorst−Pack k-points and a 350 eV planewave cutoff. The computational hydrogen electrode model was used to calculate the energy of reaction intermediates and an applied potential of 2.11 V vs RHE is applied in all energy calculations.[117] Solvation effects were modeled by reporting all energies with the corrections gained from the implicit solvation model, VASPsol.[118]

There are many variables that can potentially impact the reaction pathways of OER and MOR on these surfaces. These variables include surface coverage, the chosen intermediates for each reaction, binding sites for said intermediates, defects on the surfaces, and solvation effects.[119] Table 2 provides a pictorial representation for the binding sites we have sampled during this inquiry. Table 3 provides a symbolic depiction of the surface structure and binding sites sampled for each reaction.

Table 2 : Pictorial representations of the binding sites utilized. Red atoms indicate oxygen atoms, grey, metal atoms. Atoms highlighted in yellow and accented with the symbol "+" indicates where each intermediate is bound.

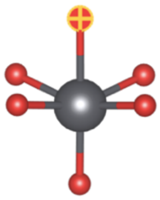
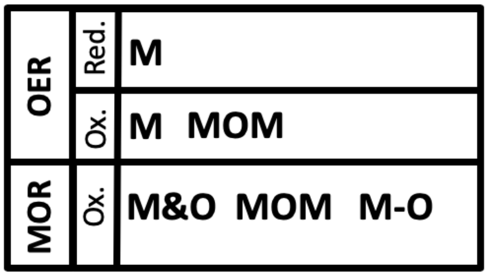
Metal Center (**M**)	Bridging Oxygen (**MOM**)	Metal with Oxygen (**M-O**)	Metal or Non-bridging Oxygen (**M&O**)
			

Table 3 : Concise symbolic representation of the binding sites and surface structures studied. Oxidized is abbreviated as Ox. Reduced is abbreviated as Red.



OER	Red.	**M**
	Ox.	**M MOM**
MOR	Ox.	**M&O MOM M-O**

9.3 Results and Discussion

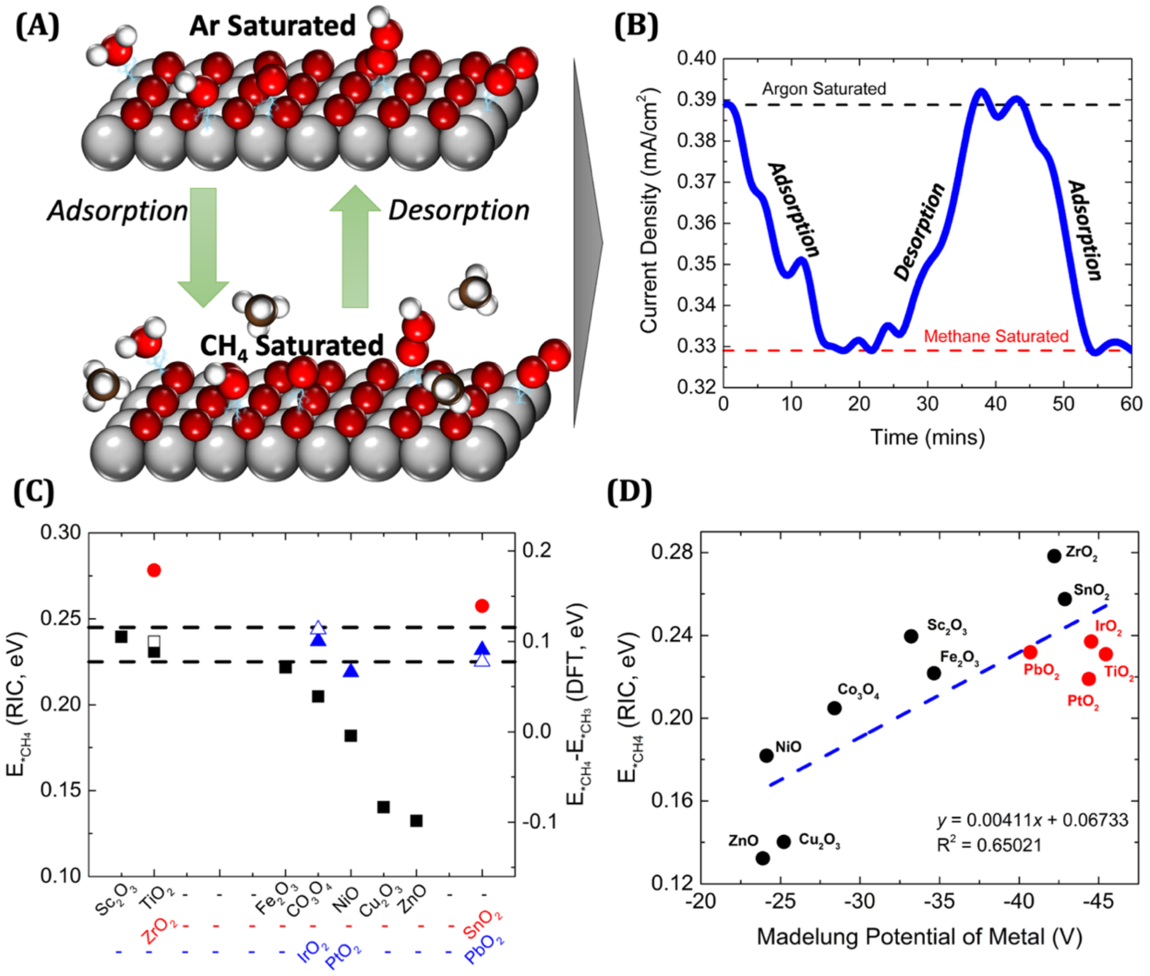


Figure 7 : (A) Schematic representation of OER intermediates on (110) surface of TMOs in Ar-saturated and CH_4-saturated electrolytes. As the Ar-saturated electrolyte is swapped with CH_4-saturated electrolyte, the CH_4 adsorbs on the TMO and reduces the number of active sites for OER. Switching the CH_4- saturated electrolyte back with Ar-saturated electrolyte allows CH_4 desorption and an increase in the number of active sites for OER. Light gray spheres are metal atoms, dark gray spheres are the oxygen atoms from the oxide layer, blue spheres represent oxygen from the electrolyte, red represents carbon, and white represents hydrogen atoms. (B) Reversible change in the OER current density with adsorption and desorption of CH_4 at a fixed applied potential of 1.52 V versus RHE. (C) Estimated binding energies of $*CH_4$ obtained using Everett isotherm for different TMOs at 1.52 V versus RHE. The MOR active TMOs—TiO_2, IrO_2, PbO_2, and PtO_2—have similar binding energy in the range 0.22 to 0.25 eV and similar activation energy in the range 0.09 to 0.12 eV. Black squares show the first row TMOs, red circles show the second-row TMOs, and the blue triangles show the third-row TMOs. (D) Scaling

relationship between measured binding energy of $*CH_4$ and the Madelung potential of metal in TMOs. The MOR active catalysts—TiO_2, IrO_2, PbO_2, and PtO_2—have a higher binding energy of $*CH_4$ and lower Madelung potential.

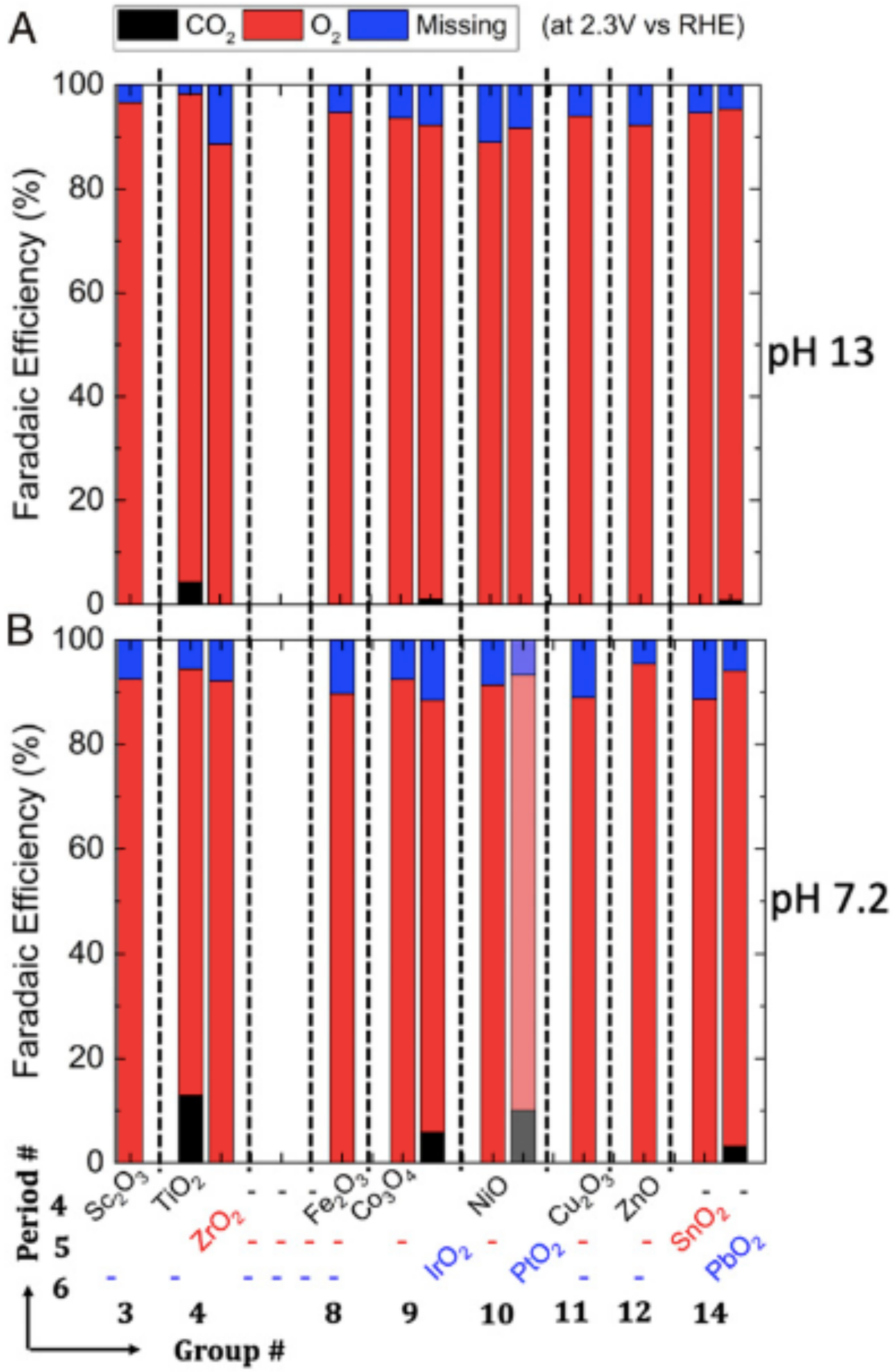


Figure 8: Product distribution and FE of OER and MOR over stable TMOs in (A) 0.1 M KOH (pH = 13) and (B) 0.1 M potassium phosphate buffer (pH = 7). PtO_2 was found to be MOR inactive in 0.1 M potassium phosphate buffer. The shaded stacked bar for PtO_2 in B shows the FE and the product distribution using 1 M KCl (pH = 7).

The MOR activity descriptors shown in Figure 7(D) indicate six catalyst candidates that could be active for MOR, namely, ZrO_2, TiO_2, IrO_2, PbO_2, PtO_2, and SnO_2. Experimental evaluation of 12 TMOs in Figure 8 shows TiO_2, IrO_2, PbO_2 and PtO_2 are the only four active catalysts for MOR. The other two catalysts - SnO_2 and ZrO_2 are not active for MOR. To further understand the activity and selectivity of MOR on the four candidate catalysts - TiO_2, IrO_2, PbO_2, and SnO_2, the OER and MOR pathways and their energy profiles were calculated using KS-DFT. Experimental conditions dictate that these TMO surfaces are oxidized, but there is little information about the nature and distribution of the oxidation states on these catalysts' surfaces. Also, it is well known that the defects on these surfaces can affect reaction pathways.[120] Therefore, we evaluate OER pathways on two extreme surface conditions- fully reduced and fully oxidized states, and then study MOR only on the fully oxidized surface. The following steps for OER were considered, where $*$ indicates the bare TMO surface:

$$* \mathrm{H_2O} \rightarrow * \mathrm{OH} + \frac{1}{2}\mathrm{H_{2(g)}}$$

$$* \mathrm{OH} \rightarrow * \mathrm{O} + \frac{1}{2}\mathrm{H_{2(g)}}$$

$$* \mathrm{O} + \mathrm{H_2O_{(g)}} \rightarrow * + \mathrm{O_{2(g)}} + \mathrm{H_{2(g)}}$$

Figure 9 (A and B) show that IrO_2 readily catalyzes OER on both the reduced and oxidized surfaces, which agrees well with previous theoretical and experimental findings.[121] However, the OER pathways on PbO_2, SnO_2, and TiO_2 surfaces are more complicated and strongly dependent on their oxidation states. The OER pathways on

TiO_2 and SnO_2 proceeds preferentially on the reduced surface at the metal center binding site, which suggests that the coordinatively unsaturated metal centers are the active sites for OER. The OER pathway on PbO_2 proceeds on the oxidized surface at the bridging oxygen binding site. Also, PbO_2 is the only surface on which it is energetically favorable for the bridging oxygen to dissociate from the surface, as seen in Figure 10. The KS-DFT-predicted OER activity on the sampled surfaces increases in the order PbO_2 (bridging O) < SnO_2 (on-top M) < TiO_2 (on-top M) < IrO_2 (on-top M), which aligns very well with the increasing order of experimental current densities for PbO_2- 0.2 mA/cm^2, SnO_2- 0.3 mA/cm^2, TiO_2- 2.5 mA/cm^2, and IrO_2- 3.2 mA/cm^2 at 2.1 V vs RHE shown in Figure 13. This indicates that the chosen computational methods can qualitatively describe the relative activities of these catalytic systems.

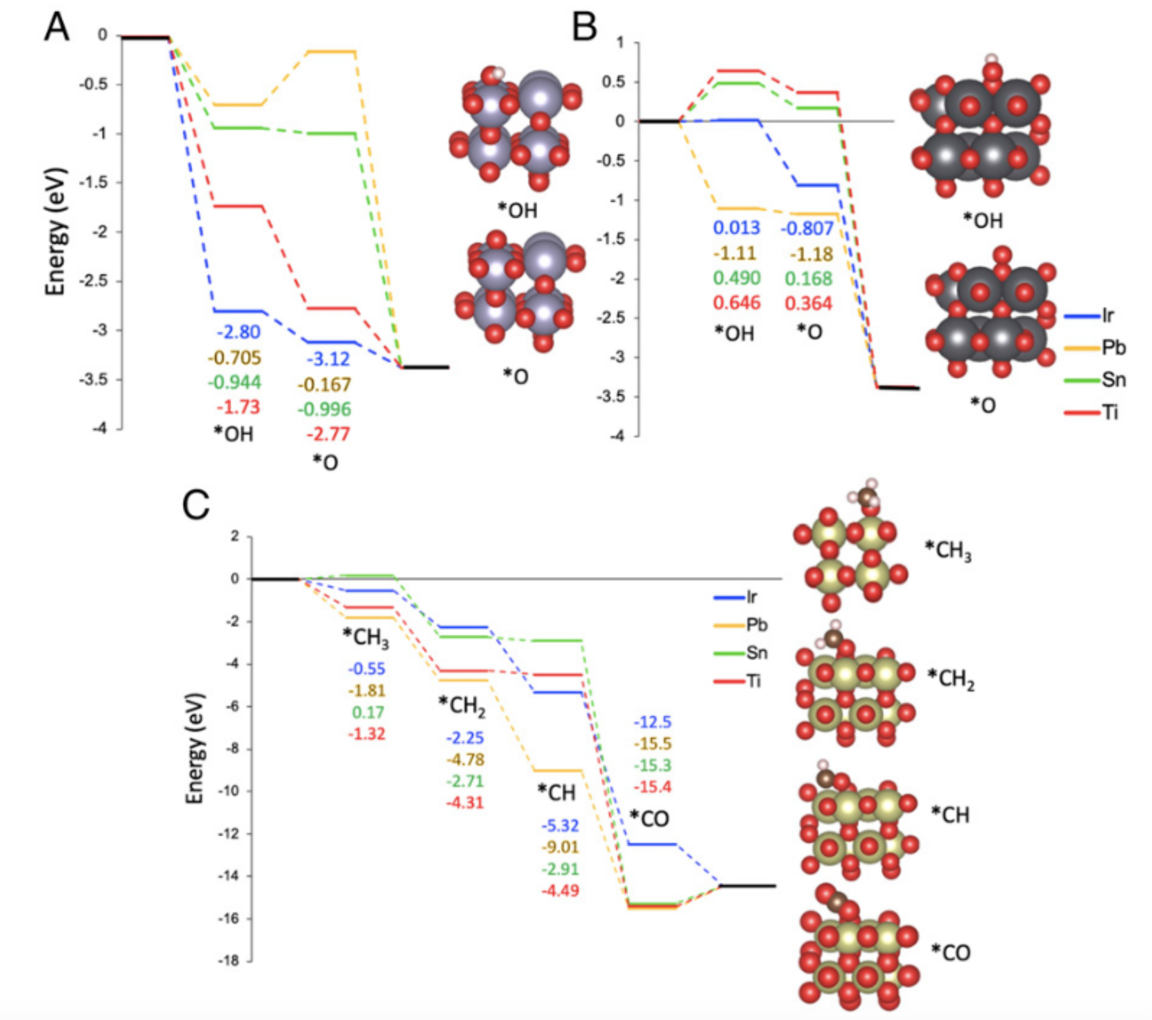


Figure 9: (A) Reaction profile for OER at the metal with oxygen binding site for the fully reduced surface (B) Reaction profile for OER at the bridging oxygen binding site for the fully oxidized surface. The relative energies of reaction are shown as a function of reaction coordinate, labeled by *X, where X indicates the adsorbed intermediate. To the right of the plot are pictorial representations of the adsorbed intermediate. Colors denote Sn (silver), Pb (gray), O (red), and H (pink). (C) Reaction profile for MOR at the metal with oxygen binding site for the fully oxidized surface. The relative energies of the reaction are shown as a function of reaction coordinate, labeled by *X, where X indicates the adsorbed intermediate along with the pictorial representations of the adsorbed intermediate. Colors denote Ir (gold), O (red), C (brown), and H (pink). All electronic energies are reported using PBE+D3 at an applied potential of 2.11 V versus RHE.

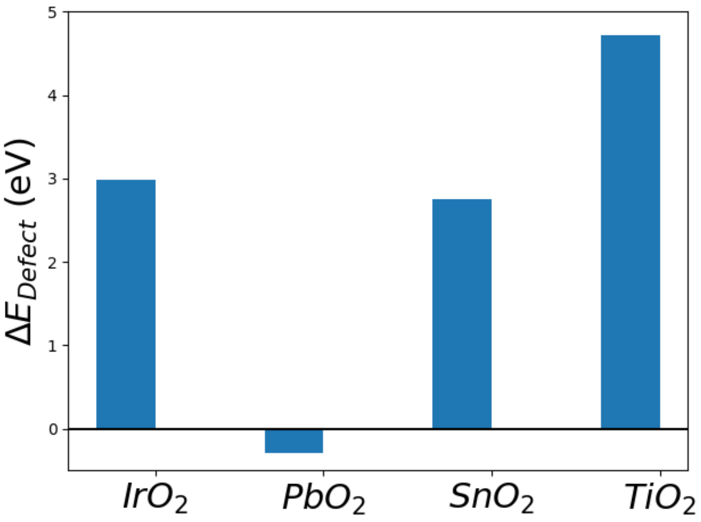


Figure 10 : The reaction energy in eV of $* \rightarrow *_{Defect} + 1/2\ O_{2(g)}$ where $*_{Defect}$ indicates the vacancy of one of the bridging oxygens on the surface.

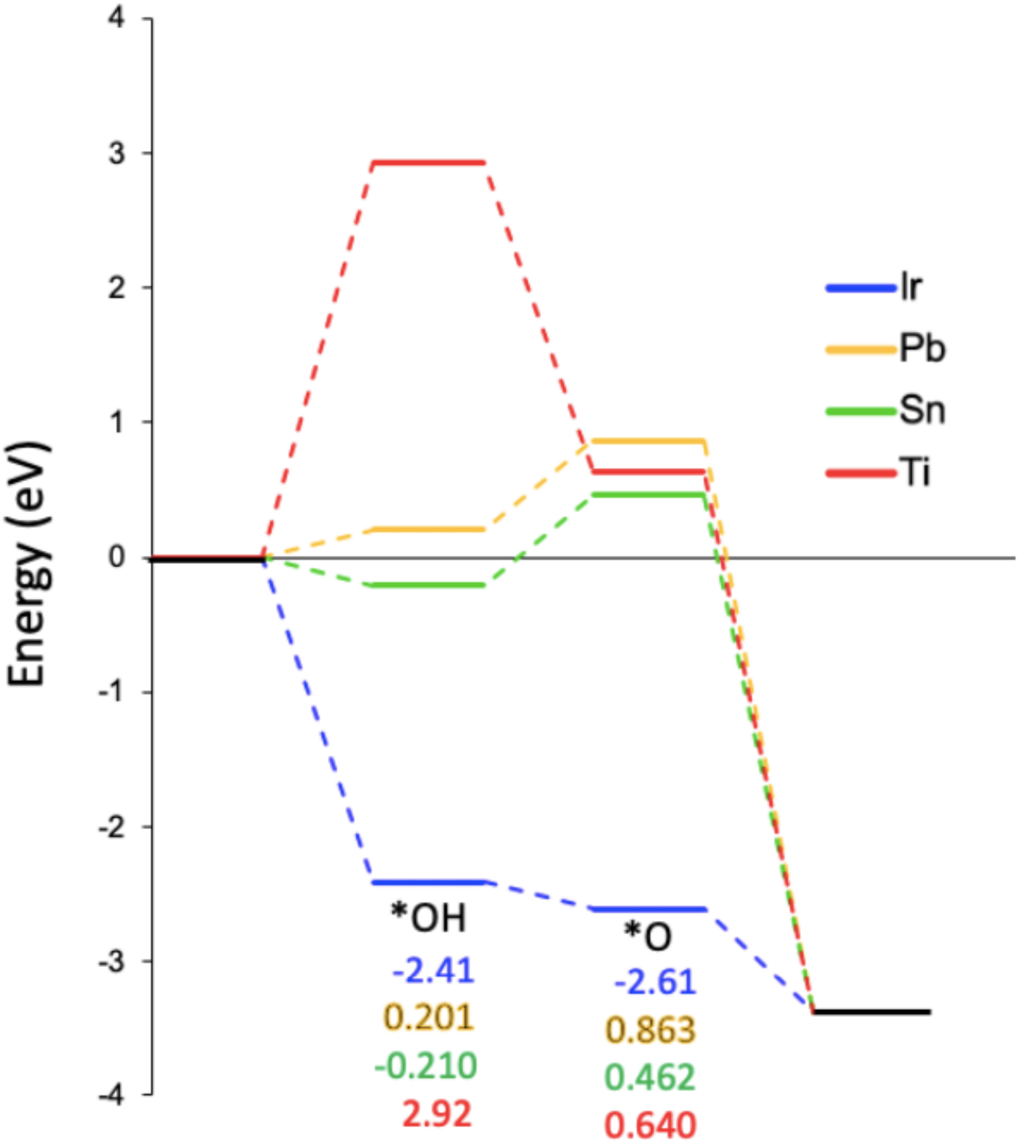


Figure 11: Reaction profile for OER at the metal center binding site on the fully oxidized surface. Shown are the relative energies of reaction as a function of reaction coordinate, labeled by *X, where X indicates the adsorbed intermediate.

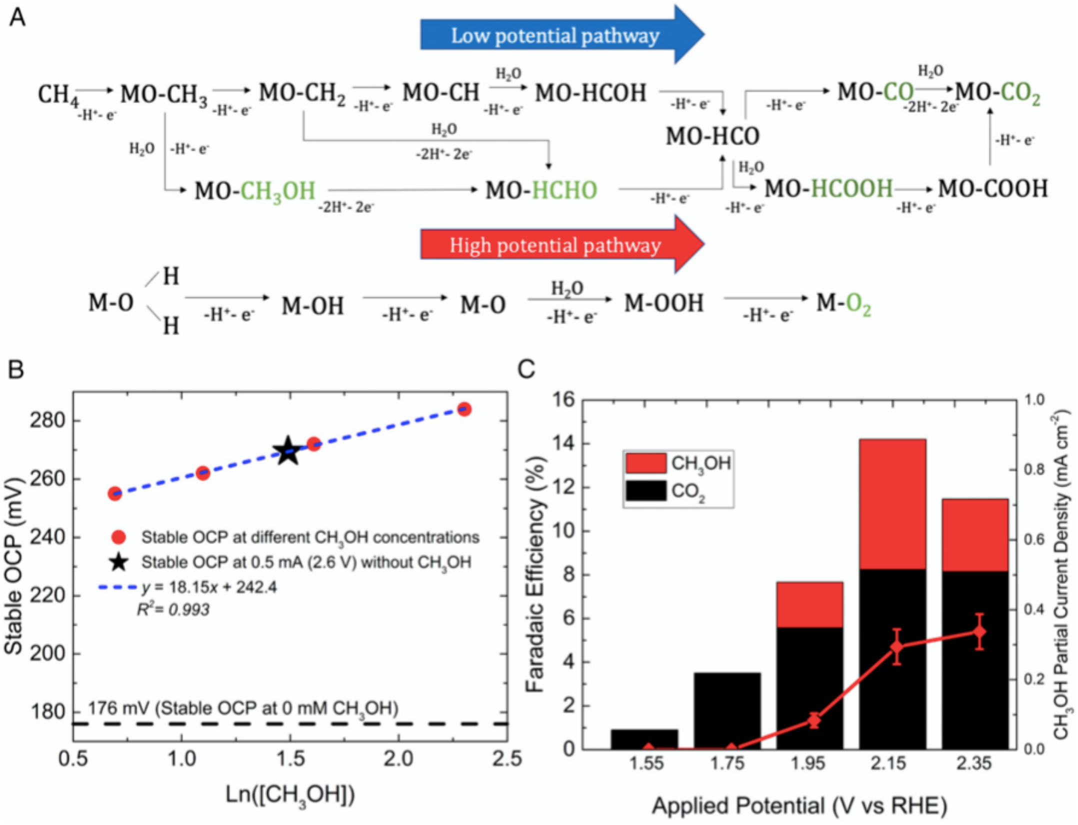


Figure 12 : (A) Possible reaction pathways for MOR and OER in an aqueous electrolyte where M represents TMO. The species in green are the possible MOR products. (B) The shift in the stable OCP with increasing concentration of CH_3OH in phosphate buffer electrolyte for an initial current of 0.25 mA at 1.8 V. The linear relation is like the Nernst equation. The increase in stable OCP from 176 mV for 0.25 mA to 269 mV for 0.5 mA of initial current in CH_4-saturated phosphate buffer electrolyte without added CH_3OH indicates the formation of $*CH_3OH$ on TiO_2. (C) FE and partial current density of CH_3OH on Cu_2O_3-TiO_2 bimetallic catalyst in CH_4-saturated, 0.1 M potassium phosphate buffer.

Figure 9 (A, B) and Figure 11 show reaction profiles for OER at various binding sites. Given this information we have concluded that OER is not readily catalyzed on the fully oxidized surface in all cases except for PbO_2. For PbO_2 we have found that the surface readily catalyzes OER at the bridging oxygen binding site, Figure 9(B), which only exists on the oxidized surface. From the information gathered, the reason for this is that on PbO_2 the bridging oxygen is easily abstracted, as seen by Figure 10. For the remainder of the metal oxide surfaces, OER is freely catalyzed on the reduced surface at the metal center binding site. Given experimental evidence, we recognize that all these metal surfaces are evolving oxygen, which we were able to show using our chosen computational methods. In this way, we were able to validate our methodology.

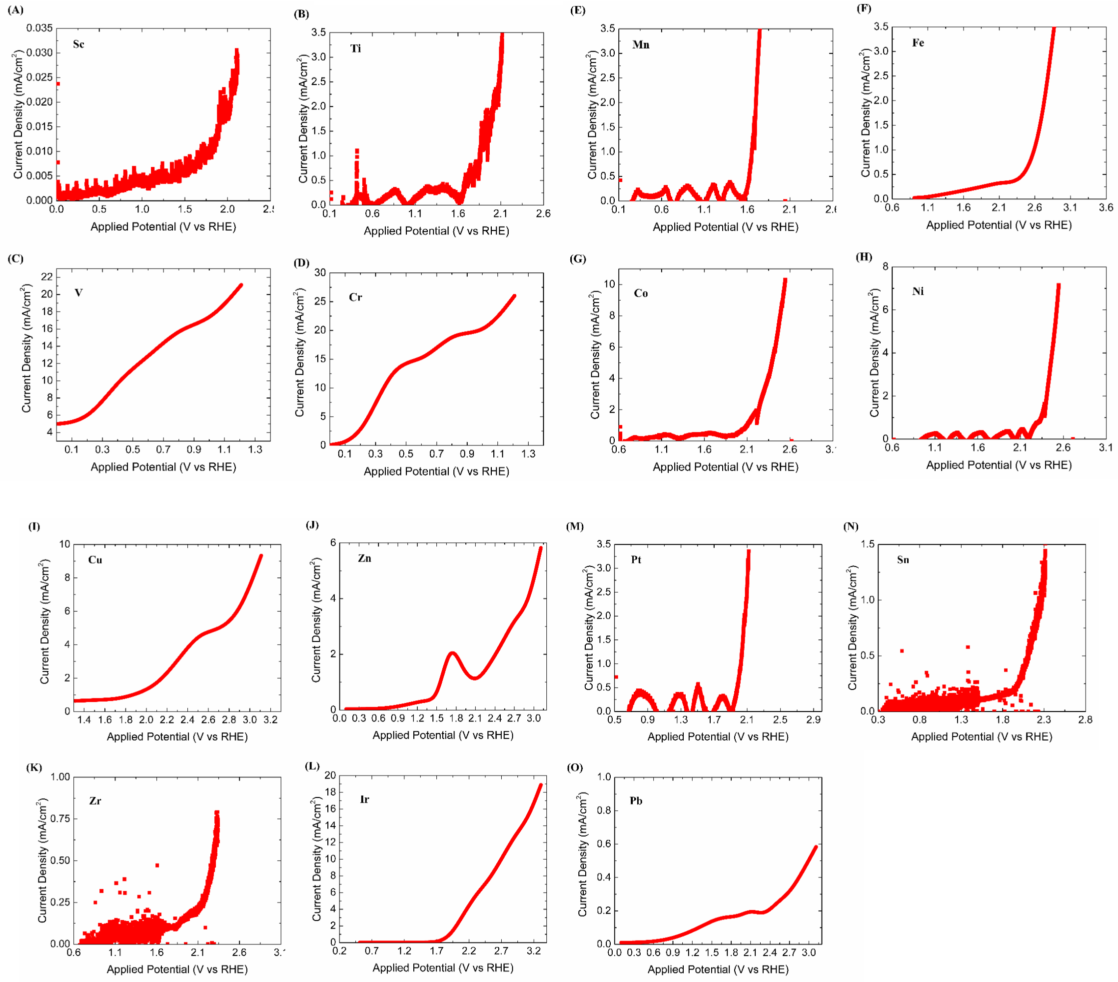


Figure 13 : LSVs of all the transition metal oxides in CH_4 saturated 0.1M phosphate buffer solution.

To study MOR on these four TMO surfaces, we considered the following steps:

$$* + CH_{4(g)} \rightarrow * CH_3 + \frac{1}{2} H_{2(g)}$$

$$* CH_3 \rightarrow * CH_2 + \frac{1}{2} H_{2(g)}$$

$$* CH_2 \rightarrow * CH + \frac{1}{2} H_{2(g)}$$

$$* CH + H_2O_{(g)} \rightarrow * CO + \frac{3}{2} H_{2(g)}$$

$$* CO + H_2O_{(g)} \rightarrow * + CO_{2(g)} + H_{2(g)}$$

where *CH to *CO is a composite step consisting of *CH → *HCOH → *HCO → *CO. Figure 9 (C) shows the first step in the catalytic cycle for MOR is comparable in energy to that of OER, which also validates E_{OH}-E_O as the common activity descriptors for OER and MOR. The reaction steps here show MOR occurring on a metal center that is occupied by an oxygen atom. This binding site for catalytic mechanisms has been discussed in zeolites and metal-organic frameworks, but to the best of our knowledge, this mechanism has had very little discussion in metal oxides.[105] The first step for MOR in Figure 9(C) is slightly higher in energy as compared to OER in Figure 9(A) and (B), which is aligned with the experimental observation of lower FEs of MOR on these TMOs. Based on the energies of the reaction pathways on different types of binding sites on TMOs, the M-O binding site is where MOR occurs favorably for all experimentally active systems. M-O is also a recognized intermediate for the OER catalytic cycle;

therefore, this binding site is readily available for catalysis on these surfaces. However, the coverage of M-O site can vary on these catalysts, as the formation of M-O site is more favorable thermodynamically on TiO_2 and IrO_2 as compared to SnO_2 and PbO_2 (see Figure 9(A)). Although the MOR is more energetically downhill on PbO_2 as seen in Figure 9(C), the lower coverage of M-O sites on this catalyst is a reason for its lower MOR activity as compared to TiO_2 (see Figure 14). The reaction steps also indicate that MOR proceeds readily on TiO_2, PbO_2, and IrO_2, but not SnO_2 as the first step in the catalytic cycle is positive for this surface. These KS-DFT results justify the observed MOR activity on TiO_2, PbO_2, and IrO_2 catalysts. To gain a further understanding of the reaction kinetics, the reaction barriers need to be calculated, which is planned for future studies.

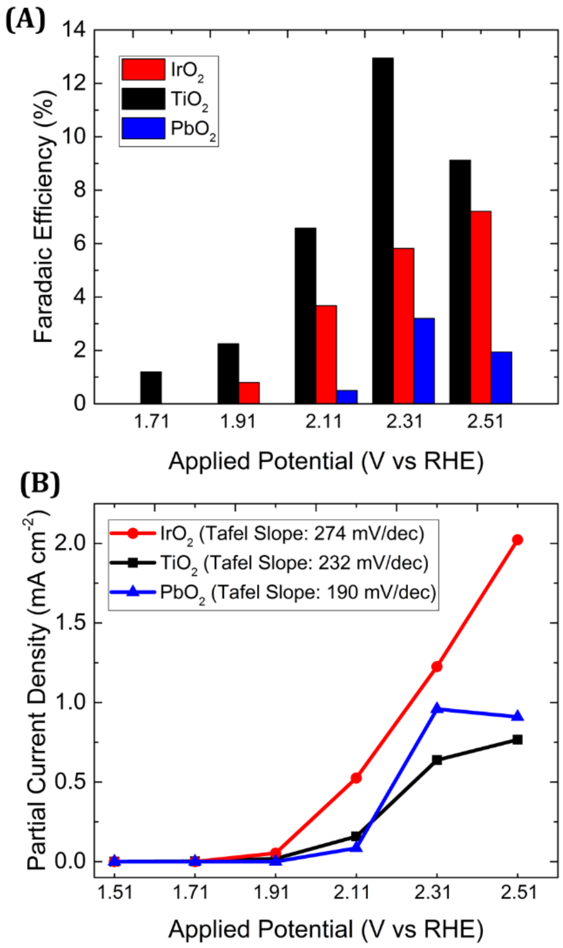


Figure 14 : (A) Faradaic efficiency and (B) partial current density of MOR producing CO_2 on TiO_2, IrO_2, and PbO_2 at different applied potentials in neutral pH, phosphate buffer electrolyte.

Figure 15, Figure 16, and Figure 17 show reaction profiles at various binding sites. Figure 17 also provides the reaction profile for an alternative mechanism beginning with methanol. For MOR, we found that during the optimization of the intermediates which were begun on a metal center often optimized to bind on a non-bridging oxygen or between the oxygen and the metal center. We also found the difference in energy between these binding sites, a metal center, or a non-bridging oxygen, to be negligible.

Therefore, we combined these two binding sites into one category, which included the version of the system that was lowest in energy, reasoning that these sites are essentially indistinguishable. This results in the reaction profile seen in Figure 15. Energetically MOR does not readily occur on any of the surfaces except SnO_2, where steps two and three are essentially the same, which warrants further study. Figure 16 shows MOR on the bridging oxygen binding site, which does not readily occur for any of the metal oxide surfaces. Figure 17 shows the reaction profile for an alternative MOR pathway that begins with methanol as listed below:

$$* + \mathrm{CH_3OH_{(g)}} \rightarrow * \mathrm{OCH_3} + \frac{1}{2}\mathrm{H_{2(g)}}$$

$$* \mathrm{OCH_3} \rightarrow * \mathrm{OCH_2} + \frac{1}{2}\mathrm{H_{2(g)}}$$

$$* \mathrm{OCH_2} \rightarrow * \mathrm{OCH} + \frac{1}{2}\mathrm{H_{2(g)}}$$

$$* \mathrm{OCH} \rightarrow * \mathrm{CO} + \frac{1}{2}\mathrm{H_{2(g)}}$$

$$* \mathrm{CO} + \mathrm{H_2O_{(g)}} \rightarrow * + \mathrm{CO_{2(g)}} + \mathrm{H_{2(g)}}$$

Here, the first step in the catalytic cycle is much higher in energy than the first step of MOR beginning with methane indicating that at the potential of 2.114 volts MOR is likely occurring with methane gas. However, it is thermodynamically favorable for MOR to occur through methanol on IrO_2 which demonstrates that the methanol pathway maybe plausible for MOR under certain conditions. In summary, we have used KS-DFT methods to study MOR and OER on a selection of metal oxide surfaces and found that OER occurs readily on these surfaces while MOR occurs on the active catalysts.

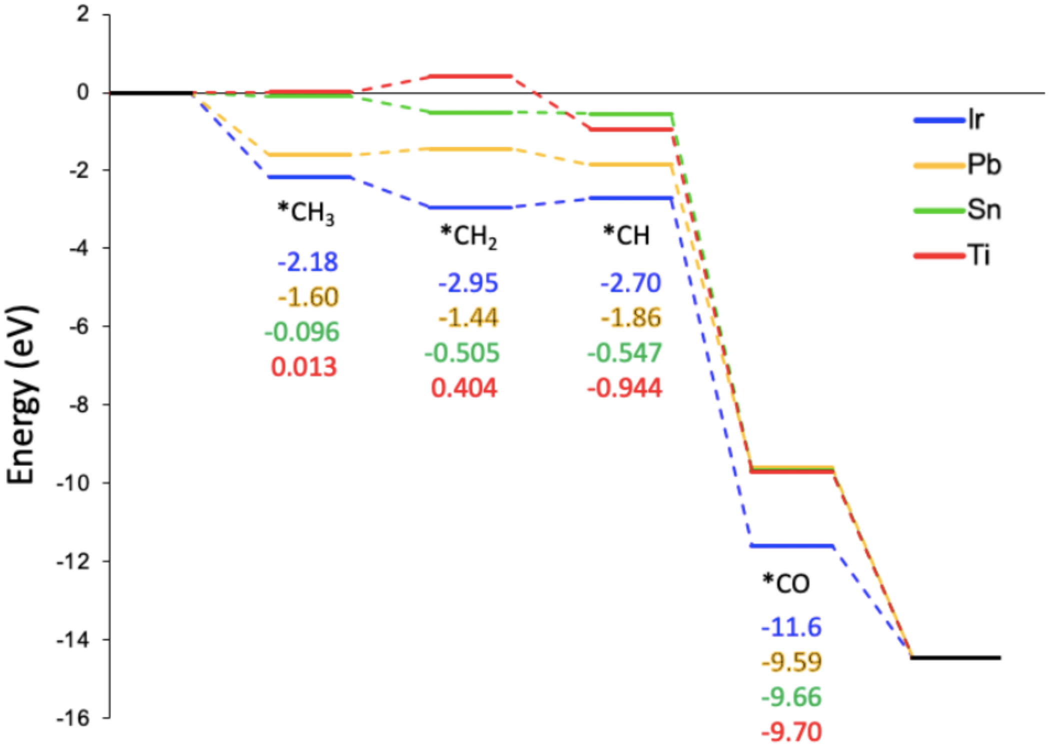


Figure 15: Reaction profile for MOR at the metal or non-bridging oxygen binding site for the fully oxidized surface. Shown are the relative energies of reaction as a function of reaction coordinate, labeled by *X, where X indicates the adsorbed intermediate.

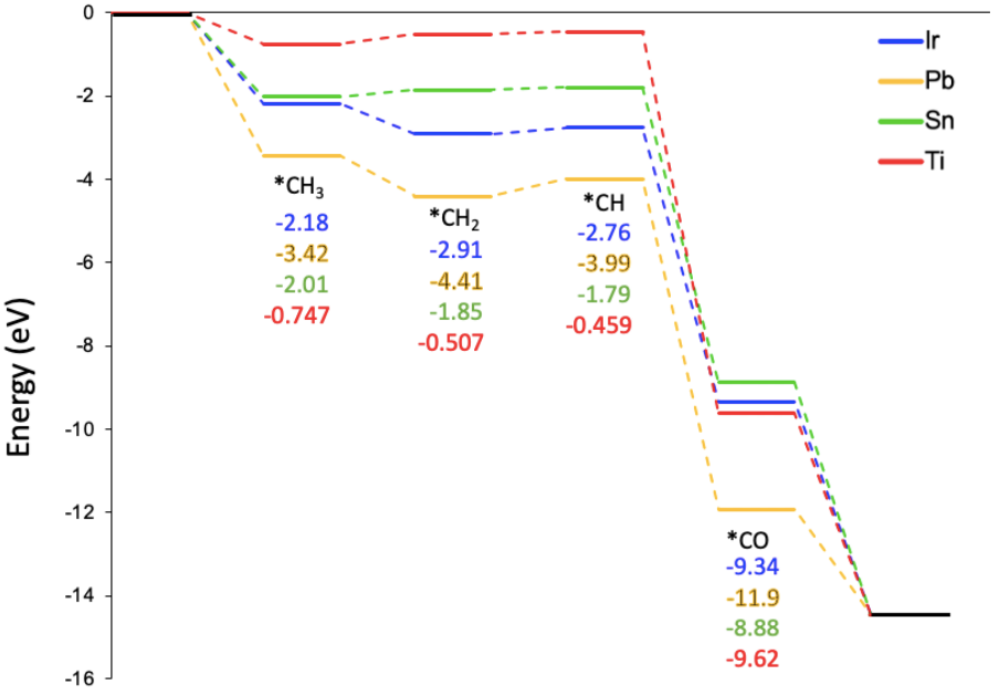


Figure 16: Reaction profile for MOR at the bridging oxygen binding site on the fully oxidized surface. Shown are the relative energies of reaction as a function of reaction coordinate, labeled by *X, where X indicates the adsorbed intermediate.

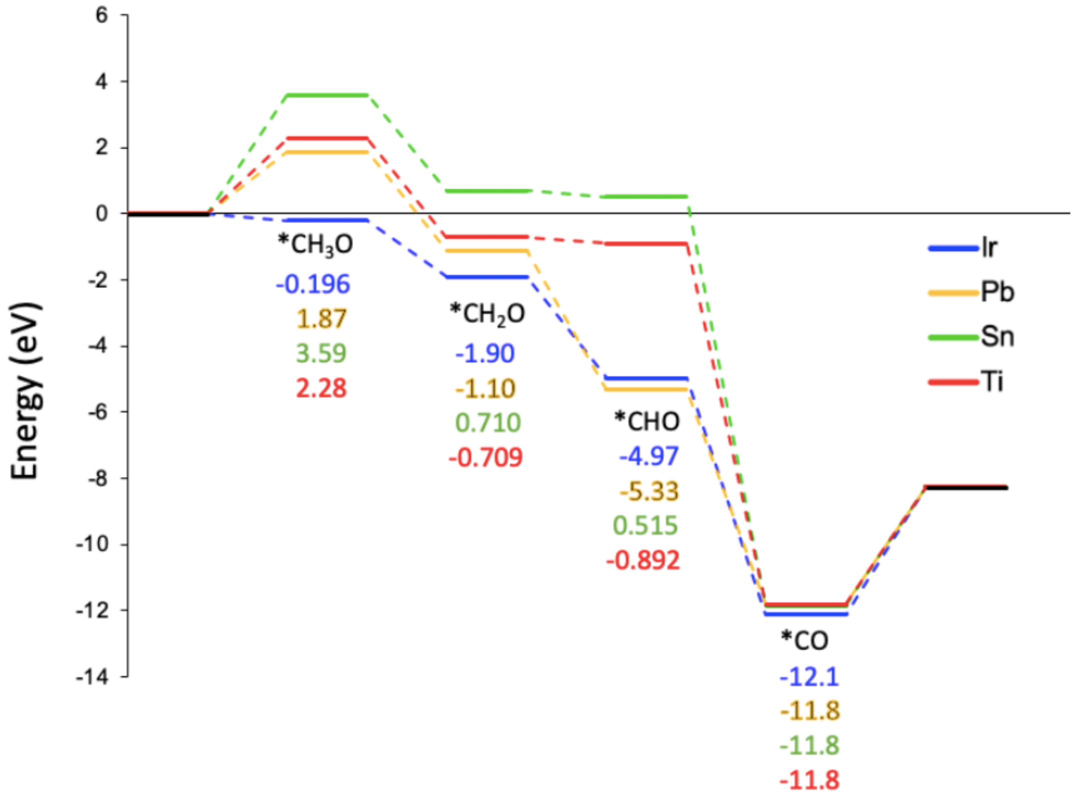


Figure 17: Reaction profile for MOR at the metal binding site on the fully oxidized surface. Here the reaction begins with methanol rather than methane. Shown are the relative energies of reaction as a function of reaction coordinate, labeled by *X, where X the adsorbed intermediate.

9.4 Conclusion

In this article, we develop a fundamental understanding of electrochemical CH_4 oxidation on TMOs and report, for the first time, the binding energy of CH_4 on TMOs under relevant conditions, activity descriptors of MOR, identification of active sites for MOR, detection of $*CH_3OH$ intermediate and its oxidation product, and strategies to harvest CH_3OH by minimizing its overoxidation.

The electrochemical oxidation of CH_4 on TMOs proceeds with the physical adsorption of CH_4 followed by the activation of C-H bonds forming either CH_x intermediates or oxygenated intermediates to yield products such as CO_2, CO, CH_3OH, or HCHO. The physical adsorption of CH_4 involves the transformation of the tetrahedral $\left(T_d\right)$ symmetry of CH_4 to a bond-angle-distorted $\left(D_{2d}\right)$structure with an H-C-H bond angle of ~120^{o}. The physical adsorption is governed by the electrostatic interaction of CH_4 with the metal of TMO. the experimental measurements show only TiO_2, IrO_2, PbO_2, and PtO_2 are the active catalysts for MOR. The inactivity of ZrO_2 and SnO_2 towards MOR is due to the poor electrical conductivity and lower population of MOR active sites, respectively.

The competitive pathways of OER and MOR are studied on the promising MOR catalysts - TiO_2, IrO_2, PbO_2, and SnO_2 using density functional theory to identify the active sites and the mechanism of the oxidation reactions. While the OER pathway on SnO_2, TiO_2, and IrO_2 is favorable on the under-coordinated metal site, this reaction on PbO_2 occurs on the bridging O site. The KS-DFT-predicted OER activity on the sampled surfaces increases in the order PbO_2 (bridging O) < SnO_2 (on-top M) < TiO_2 (on-top M) < IrO_2 (on-top M), which aligns very well with the increasing order of experimental current densities. The M-O (metal-oxygen) intermediate formed on these TMOs is also the most favorable active site for MOR. The energy profiles of MOR on M-O sites are energetically favorable on these TMOs, except SnO_2, for which the first dehydrogenation step is positive. The lower OER current and hence lower M-O active site coverages on SnO_2, along with the higher energy barrier for MOR, are the reasons for its inactivity towards MOR.

10 Computational Chemistry for Metal–Organic Frameworks: Machine Learning Focused Methods

This chapter describes the outcome of a collaborative research project carried out by the authors of the following:

Indrani Choudhuri, Jungyun Ye, Daniel S. Graham, Brianna Collins, Jason D. Goodpaster, and Donald G. Truhlar

Reported is work conducted directly by and in consultation with the author.

10.1 Introduction

Metal–organic frameworks (MOFs) are networks, usually three-dimensional, of inorganometallic nodes connected by organic linkers. Some involve coordinated metal ions and may also be called porous coordination polymers. They have attracted widespread research interest due to their ultrahigh porosity, huge surface area, remarkable stability, and crystalline characteristics. These unique and extraordinary structural properties of MOFs enable a wide range of applications, particularly in the field of heterogeneous catalysis,[122,123,124,125] photocatalysis,[126,127,128] electrocatalysis,[129,130] chemical adsorption,[131] separation,[132] and magnetism.[133] Most of these applications depend on the availability of the void areas (pores) of the MOFs. MOFs generally consist of metal nodes bridged by organic linkers[134,135,136] and the topology of MOF pores is mainly determined by the composition of nodes and linkers.[137,136]

Due to the simplicity of the MOF framework and the high number of possible nodes and linkers, there are theoretically millions of potential MOFs. Here, we can address the importance of theory and computational methods and talk about the synergistic effects of computational methods and techniques. Newly emerging machine learning techniques provide a promising pathway for significant discoveries within metal-organic frameworks (MOFs). Machine learning (ML) has a long history of use for discovery in science. For example, within materials science, ML has been used for numerous applications, including the optimization of organic semiconductors, metal oxides, and conductive 2D materials.[138,139,140,141,142] As for MOFs, ML techniques can be used to uncover hidden trends in data that may not be obvious yet are crucial in understanding the material performance. Moreover, ML can identify possible structures that would show the best performance for applications of interest that are hidden in the vast number of imaginable MOFs, making it essential to understand how ML methods have been applied for MOFs to foresee fruitful directions for this technique.

10.2 Computational Methods

ML is a branch of computer science where models are trained to analyze given data to learn trends and patterns that allow the models to provide accurate predictions. ML models for MOFs are as diverse as MOFs; no two models are alike, each tailored by the available information and the desired outcome. ML models for metal-organic frameworks, as discussed below, loosely fall into the following categories: (i) the screening of databases for MOF materials with desired properties and (ii) the screening of data sets for hidden relationships between the inherent qualities of the MOF and its properties, with a generous overlap between the categories.[143]

Here, we lightly review a subset of potential ML algorithms used in the applications below. Neural networks (NNs) are a set of branching algorithms that are designed to recognize patterns. Neural network potentials (NNPs) are arising as particularly interesting in the context of molecular dynamics. NNPs utilize NNs in statistical learning models that approximate the potential energy of molecular systems and can be used essentially like force fields, representing the level of theory used to train the NNP. This enables the calculation of DFT or higher levels of information at MM costs.[144] One of the most used ML techniques in chemistry is kernel ridge regression (KRR). Kernel methods, such as KRR, move beyond linear methods by transforming into higher-dimensional space, where nonlinearities in data can be understood efficiently. Support vector machines (SVMs) are equivalent to KKR methods except for minor loss function differences, which is how the ML algorithm minimizes error. Along this same line of ML, techniques are the more standard linear and logistic regression methods. Gaussian process regression (GPR) is another commonly used technique and is directly in contrast to standard regression methods; here, one does not use a specified basis function but allows for a family of different possible parts which are then tailored toward data.[145] Decision Tree (DT) algorithms are strikingly different from the above-stated methods but have nevertheless found use in current research concerning metal-organic frameworks. DT algorithms are flowchart-like models built by splitting the data based on values of conditional rules. As more than one or two conditional statements are usually needed to describe a data set, decision tree algorithms are prone to becoming over-fit, as they learn highly irregular and not reproducible patterns in the data. Usually multiple decision trees are employed, which is referred to as the random forest (RF) method.[146] For more information on the details of ML techniques in porous materials, please see the excellent review by Jablonka et al.[147]

10.3 Properties & Applications of MOFs

MOFs are popular for their wide range of applications in the fields of electronics, magnetism, gas storage, adsorption and separations, water filtrations, among many others. In this review we will give an overview of some of the major applications which are extensively studied.

10.3.1 Structural Properties

The link between the structure and properties of materials has been thoroughly explored using ML methods. For instance, Moghadam and coworkers investigated the mechanical properties of MOFs using a multiscale modeling approach.[140] They used high-throughput molecular simulations for 3,385 MOFs with diverse network topologies. They then developed an NN-based machine-learning algorithm to predict the mechanical properties of MOFs automatically. This allowed them to identify causal relationships between the mechanical robustness of MOFs in equilibrium, as expressed by the bulk and shear moduli and the chemical-structural properties. Moghadam and coworkers went on to perform *ab initio*-based force field molecular dynamics calculations to determine the effects of temperature and pressure on mechanical robustness. High-throughput screening of mechanical properties in MOFs provides in-depth insights into structure-mechanical stability relationships. It enables the rational design of MOFs with better mechanical properties.[140] This is key to bringing the application of MOFs closer to reality and highlights the undeniable relationship between the structure and mechanical properties of MOFs.

10.3.2 Electronic and magnetic properties

The electronic properties of different MOFs have become a front-line area of research due to their potential application in developing semiconductor and supercapacitor technologies, high surface area conductors, thermoelectrics, and photocatalysts. Semiconducting properties are typically defined by the energy gap or barrier (or band gap) between the conduction band, or highly occupied crystal orbital (HOCO), and valence band, or lowest unoccupied crystal orbitals (LUCO). The metal oxide nodes are primarily responsible for the semiconducting behavior of the MOFs,[148] while the delocalized p-bonding networks of the organic linkers provide high conductivity. These two unique electronic properties make MOFs desirable in semiconducting and superconducting technology.

Conductive MOFs are intriguing for electrical applications as they maximize the contact between the conductive pathways and molecules of interest. Narayan and coworkers showed that MOFs could be synthesized to become electrically conductive while preserving their porosity.[149] Marinescu and coworkers identified the first experimentally observed MOF that exhibits band-like metallic conductivity.[150] Here, ML techniques can be incorporated to aid in discovering novel conductive MOFs. He et al. revealed that ML methods are feasible for predicting metallicity in MOFs. They used supervised classification learning to identify metallic MOFs in the CoRE-MOFs database and were able to identify 6 MOF crystal structures that are metallic at the level of semi-local DFT band theory.[151]

10.3.3 Thermal Properties

The thermal expansion coefficient is defined as the fractional change in size per unit change in temperature (in degrees) of a material at some constant pressure. MOF-5 is known to show negative thermal expansion, which means that it shrinks as the temperature rises. Zhou et al. analyzed the origin of the large negative thermal expansion s in MOF-5 using lattice dynamic calculations.[152] Accordingly, their phonon softening behavior arises from the low-frequency lattice phonon modes, which are responsible for negative thermal expansion.[153,154] The IRMOF family and HKUST-1 also shows negative thermal expansion, which was explored with the help of *ab initio* study.[155,156,157] MOFs which exhibit this behavior tend to do so due to the thermal motions of their organic linkers.[158] Within the realm of ML, negative thermal expansion behavior can be challenging to capture owing to the lack of relevant data for this phenomenon. However, Marco Eckhoff and Jörg Behler was able to capture this behavior through the development of first principles quality reactive atomistic potentials, which were based on molecular fragments of MOF-5.[152] They were able to convincingly establish that even in the absence of bulk structures in the construction of the NNP, all the investigated bulk properties, such as lattice parameters, elastic constants, bulk modulus, and NTE, are accurately predicted.[154]

10.3.4 Photocatalysis

The energy bandgap of materials regulates the absorption of photons for photocatalytic materials. Here, ML methods can be instrumental in their ability to screen databases for a particular property, namely the bandgap. One such study that displays the utility of ML methods for this endeavor is the work of Lee et al., who constructed an SVM-based ML model that was able to predict bandgaps with an RMSE of 0.24 eV over 270 inorganic molecules.[161] However, to calculate accurate band gaps, one needs to employ the use of exact and high-cost methods. Recently, ML methods using Co-kriging regression have been developed to take in low-fidelity data and, through careful training, reproduce higher-fidelity data. One such study is from Pilania et al., who were able to construct a model to predict the band gap of material at the HSE06 level based on the corresponding PBE band gap.[162] While their methods were not directly employed for MOF materials. One can easily imagine how this methodology can be applied to MOFs in a future study.

10.3.5 Electrocatalysis

Electrocatalytic reduction is a promising pathway for the sustainable conversion of CO_2 into fuel and energy products. Nørskov and coworkers proposed a theory guiding the chemisorption of reaction adsorbates in terms of the electronic structure of the materials, which is called the d-band theory.[163] Noh and coworkers took advantage of this theory and constructed ML models using KRR and NNs that can be used to predict the binding energy of surface adsorbates using a linear muffin-tin orbital theory based d-band width. Interestingly, even though they did not use any *ab initio* data in their model, they achieved very accurate results in predicting *CO binding energy.[164] Noh et al. then screened a set of alloy catalysts for the CO_2 electrochemical reduction reaction by estimating their *CO binding energies, thereby displaying the applicability of their model, which exceeds their initial test sets. Overall, Noh et al. showed that being able to expeditiously estimate *CO binding energies using the easy-to-compute input features is extremely exciting for the screening of materials for their ability to reduce CO_2 readily. Recently, Mao et al. developed an approach using similar descriptors and regression techniques to screen MOF materials for those that will readily reduce CO_2.[165] Interestingly, they found that the limiting potentials of MN_4-MOFs (M = Ti to Cu), which were determined by the formation energy of *OCHOH and *OCH_2OH species, are related to the d-band center of the material. Overall, Mao et al. found that FeN_4-MOF was the most active catalysis in their study of the reduction of CO_2, with a low limiting potential of only −0.41 V. This limiting potential is very competitive with other reported catalysis for CO_2 reduction, highlighting the importance of ML techniques in this area.

10.3.6 Gas adsorption/separation/storage

Generally, one leverages differences among adsorbates, such as the rate of transport, to make the phenomena of gas separation and storage possible. ML techniques have thrived in this area, particularly concerning high throughput screening of databases to find materials that are best suited for these processes.[166] Fernandez and coworkers constructed several seminal regression models using radial distribution functions as predictor variables to predict CO_2 and N_2 uptake.[167] Fernandez et al. went on to include more structural variables, like void fraction and pore size, to predict CH_4 uptake.[168] They also applied a classification approach based on quantitative structure−property relationships to predict top-performing MOFs for CO_2 adsorption with 94.5% accuracy.[169] More recently, Gharagheizi and coworkers used ML algorithms to predict the adsorption of mixtures in MOFs in order to distinguish structures that may separate azeotropic mixtures.[170] In order to achieve this, they used adsorption isotherm data with a multiple regression genetic algorithms to successfully develop a model based on molecular descriptors.[171,172] These experiments highlight the utility of ML methods for gas storage and separation.

10.4 Conclusion

Recently, ML has been developed to uncover correlations in data sets. While significant and trailblazing progress has been made thus far, the application of ML methods to MOFs has only just begun. We foresee a substantial escalation in scientific studies dedicated to developing and expanding ML techniques to establish and predict the material properties of MOFs shortly.

11 Predicting Adsorption Energies on Bimetallic alloys: Examining Machine Learning models and Categorical embeddings

Brianna A. Collins & Jason D. Goodpaster. *In preparation*.

11.1 Introduction

First-principle methods are crucial in the discovery and development of new catalytic materials. Further, reaction barriers in catalytic processes are often estimated using adsorption energies of reaction intermediates as related in the Brønsted–Evans–Polanyi equation. Thus, adsorption energies are useful in the description of catalytic activity.[173] As many computational studies have shown,[174,175,176] there exist linear relationships between the adsorption energies among reaction intermediates which significantly influences the concept of large-scale screening of catalysts for an optimal species. Abild-Pederden et al. showed that there exists linear correlations between the binding energies of CH_x*, NH_x*, OH_x*, and SH_x* with those of C*, N*, O*, and S*, respectively.[113] Taken with the Brønsted–Evans–Polanyi equation, the reaction energy and activation energy can be correlated with the adsorption energy of one of the intermediates. The expression of catalytic activity as a function of the adsorption energy is normally depicted by Sabatier volcano plots.[117] These linear scaling relations for surface intermediates are often utilized in the development of predictive models and typically use easy to compute descriptors. Thus, one can use the predicted adsorption energy and the related volcano plot to enable the calculation of the catalytic activity of a surface binding site, enabling the large-scale screening of catalysts.

Traditionally, the adsorption energies are calculated by *ab initio* methods, like Kohn–Sham density functional theory (KS-DFT).[9] The KS-DFT calculations of adsorption energies on surfaces often demand large unit cells which are prohibitively expensive to do on the large scale. Thus, it is necessary to develop alternative models that can predict adsorption energies on the surfaces of complex systems. Many groups have taken this as motivation to make use of advances in data science and machine learning (ML) and have had success. Toyao et al. developed an effective ML model for predicting the adsorption energies of CH_4 related species on the Cu-based alloys with their leading model having an average root mean squared error below 0.3 eV.[177] Noh et al. used non-*ab initio* linear muffin tin orbital-based input features to predict the *CO binding energies on alloy systems and obtained a root mean square error of 0.05 eV.[164] Fung et al. demonstrated an approach to predict adsorption energies using a convolutional neural network to automatically gather features from the electronic density of states yielding a mean absolute error of around 0.1 eV.[178] Please see the excellent reviews by Chen et al.[179], Guan et al.[180], Lamoureux et al.[181], and Xu et al.[182] for a continued exploration of this topic.

This work aims to further the field of computational catalyst screening by showing that nominal categorical data, such as the chemical composition of a surface, can be useful in constructing said regression models. Here we postulate that readily available data, such as material properties like density, along with the stated categorical information, are sufficient to build a generalizable ML model of adequate accuracy. To do this, we focused on building a ML regression model to predict the adsorption energy of twelve diverse adsorbates on bimetallic alloys, utilizing the dataset by Mamun et al. to provide the adsorption energies, as well as information regarding the surface composition, the identification of the adsorbate, and information about the adsorption site.[183] In addition, we also used information readily available on NIST's search engine,[184] such as thermodynamic information concerning the adsorbates, as well as physical property information from the materials project database.[185] Overall, this led to the acquisition of a plethora of categorical information, such as the type of site the adsorbate was bound (e.g. “hollow” site). Typically, this information needs to be either transformed or disregarded to perform a ML regression task on a dataset. Here we employed entity embedding, a technique from the natural language processing field, to fully utilize our categorical information. This method maps categorical variables as a function approximation problem in Euclidean space which is learned by a shallow neural network using a standard supervised training process yielding the entity embeddings (i.e. features).[186] Using entity embedding is advantageous because mapping similar values close to each other in the embedding space reveals the intrinsic properties of the categorical variables, which can be leveraged in subsequent regression tasks. After transforming the data, we built a regression model using the Catboost[187] regression method, which yielded a 2.35 kcal/mol mean absolute error on the reserved test set.

11.2 Computational Details

11.2.1 Development of features

In the field of catalysis, features provide correlations between fundamental properties, such as the electronic structure, and the target, here the adsorption energy. The use of easily accessible features has increased the discovery of new catalytic materials. First-principles-based features, as pioneered by Nørskov and colleagues, who demonstrated how metal electronic structure features, such as the metal d-band center, correlate with the binding energy of surface, are very popular in the literature although their use comes with disadvantages.[188] Several groups have pioneered the use of non-*ab initio* features as an effort to reduce the computational cost associated with developing higher level features. Examples include geometric and atomic features of the potential active site[189]; generalized coordination descriptors, such as typical valence, number of free bonds, and ionic radius have are also popular.[164] Some have proposed the use of easy to compute linear muffin-tin orbital based features such as the d-band width and electronegativity.[164] Gusarov et al. developed features based on the Fukui function and its projection onto the Connolly surface formed by the smoothed van der Waals spheres of individual atoms in an attempt to capture the topology of a surface and analyzed the correlation between the proposed descriptor and the surface adsorption energy of CO.[190] Esterhuizen et al. used generalized additive models (GAM) to predict small-molecule adsorption energies on Pt alloys using non-KS-DFT calculated metal properties such as the pseudopotential radius, the number of valence electrons, the electronegativity and the molar volumes.[191] Here our goal was to utilize simple and accessible features from NIST,[184] Materials Project,[185] and Mamun et al.[183] in the hope that this would allow us to create a model that is both interpretable and easy to construct.

Data preprocessing can greatly influence the outcomes of the model and the ability to interpret the results. Appropriate preprocessing is pivotal to the accuracy and relevance of the ML model. Given that ML models are based on mathematical equations, it is therefore problematic if categorical information such as “hollow” or “hcp” are desired to be used in the algorithm. This information must be converted into machine-readable numerical values before it can be utilized. One-hot encoding is one commonly used method of transforming categorical data.[192] With one-hot, each categorical value is converted into a new feature and assigned a binary value of 1 or 0 to those columns. Each integer value is represented as a binary vector. This procedure generates several new variables and thus using one-hot can cause difficulties if the original column has many unique values. Another disadvantage of one-hot encoding is that it produces multicollinearity among the various variables, lowering the accuracy of a model constructed using data preprocessed as such.[192]

Entity embedding is another method used to transform categorical variables into usable information for a ML task.[186] Entity embedding involves encoding categorical information and then using a shallow neural network to perform a regression between the encoded information and the target. The features extracted from this process are the weights optimized during the regression. Therefore, the features contain information about how each category is related to the target. The process of entity embedding begins by mapping each state of a discrete variable to a vector as:

$$e_i : x_i \longmapsto \mathbf{x_i}$$

$\mathbf{x_i}$ is then label-encoded as:

$$u_i : \mathbf{x_i} \longmapsto \boldsymbol{\delta}_{x_i\alpha}$$

Where $\boldsymbol{\delta}_{x_i\alpha}$ is Kronecker delta and the possible values for α are the same as $\mathbf{x_i}$. If m_i is the number of values for the categorical variable $\mathbf{x_i}$, then $\boldsymbol{\delta}_{x_i\alpha}$ is a vector of length m_i, where the element is only non-zero when $\alpha = \mathbf{x_i}$. The output of the extra layer of linear neurons given the input x_i is defined as:

$$\mathbf{x_i} \equiv \sum_{\alpha} w_{\alpha\beta} \boldsymbol{\delta}_{\boldsymbol{x_i \alpha}} = \boldsymbol{w}_{\boldsymbol{x_i \beta}}$$

Where $w_{\alpha\beta}$ is the weight connecting the label encoding layer to the embedding layer and β is the index of the embedding layer. The mapped embeddings are just the weights of this layer and can be learned in the same way as the parameters of other neural network layers. All embedding layers and the input of other continuous variables are then concatenated. The network can thus be trained with the standard back-propagation method.[186] In this way, the entity embedding layer learns the intrinsic properties of each category, while the deeper layers form complex combinations of them.

To ascertain which method of preprocessing was most appropriate for our data set we used both methods and compared correlations between the features and the target. We also compared the accuracy of several models built with each data set. In Figure 18, we have plotted the Pearson correlation coefficients for all features and the target in a heatmap with the target column shown separately for clarity for both the dataset with the nominal information one-hot encoded and entity embedded. From Figure 18 the data that was entity-embedded is significantly more correlated to the target than the data that was one-hot encoded. Moreover, this would indicate that a model built using entity embedding should perform better than that of one built using the dataset that had been one-hot encoded. Therefore, we tested this hypothesis by constructing five different tree-based machine learning models using data that had been one-hot encoded and entity embedded and compared the resultant metrics as shown in Figure 19. All models were left with default hyperparameters, and each metric shown is the average of 10-fold cross validation with the standard deviation shown as the error bar. Here only tree-based models were sampled because they are invariant to the scaling of the data and robust to outliers, which were not screened from the dataset. Tree-based methods also contain inherent feature selection, which is taken advantage of in this study.[193] Overall, the Catboost regression algorithm performs best in comparison to all tested ML algorithms (Xgboost,[194] Gradient Boosting,[195] Lightgbm,[196] ExtraTrees[195]). Models constructed using data that had been embedded outperformed on all metrics and in all models thus clearly demonstrating the importance of preprocessing and the intelligent inclusion of nominal information in building regression models.

11.2.2 Development of Model

To train our ML model, we used a dataset containing 52,792 adsorption energies on 1088 unique bimetallic alloy surfaces.[183] These surfaces are composed of 37 metal elements in varying stoichiometric ratios. Adsorbates included in this set are the monatomic adsorbates H, C, N, O, S, and polyatomic counterparts, CH, CH_2, CH_3, NH, OH, SH. The adsorption energies range from at least -25.64 eV to 20.52 eV which implies ample sampling of both weak and strong bonds in adsorption systems. Therefore, it is assumed that this dataset represents a subset of important intermediates adsorbed on surfaces exhibiting an extensive range of chemical environments and can be used to develop a model that is generalizable. The type of model, CatBoost[187], was chosen based on the results discussed above and shown in Figure 19. Hyperparameters were tuned using Optuna[197] search, where we were careful to control for data leakage.[198,199]

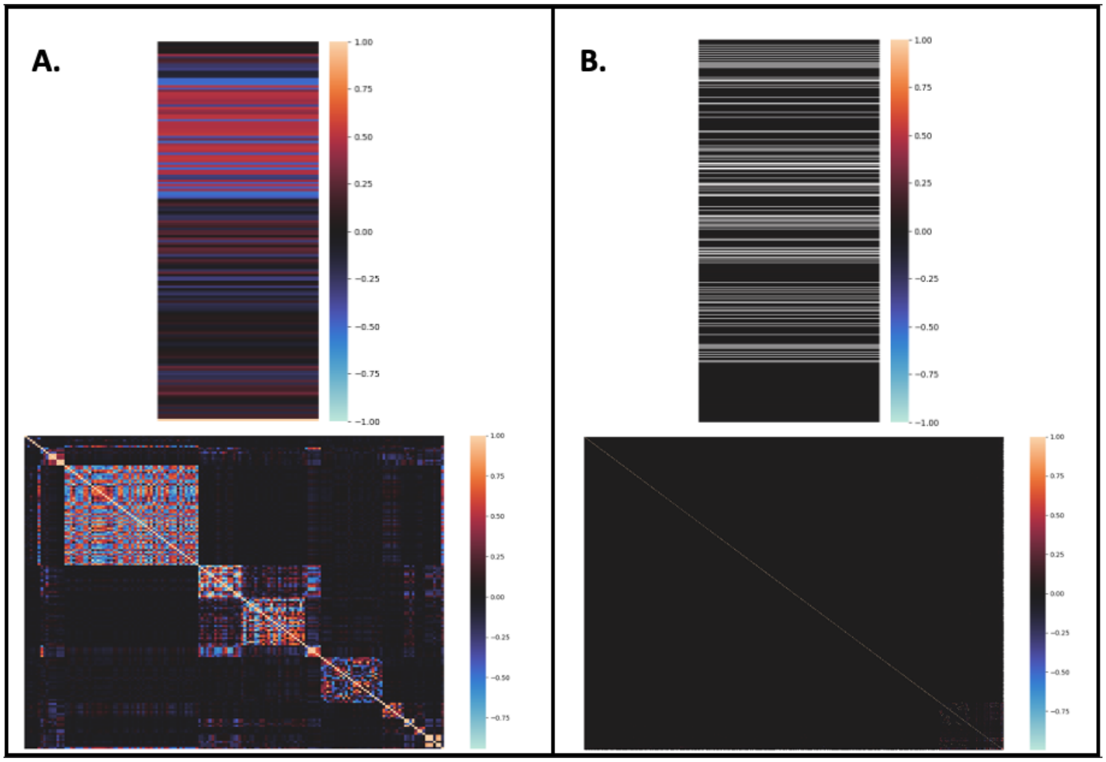


Figure 18 : Pearson correlation coefficients plotted as a heatmap with the top image as a larger representation of the coefficients in the target column. A indicates that the data shown has been entity embedded, B indicates that the data had been one-hot encoded.

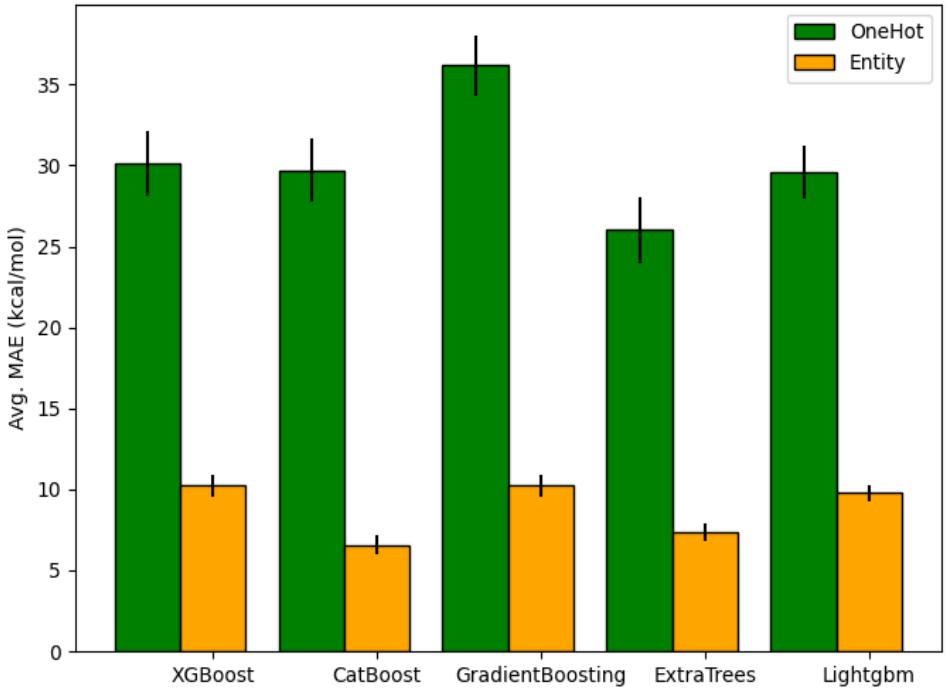


Figure 19 : Average MAE over 10 cross-validation attempts for various regressors. The data has been preprocessed as one-hot encoded and entity embedded as noted.

11.3 Results and Discussion

11.3.1 Evaluation of ML performance

For ML models tasked with regression, the mean absolute error (MAE) and root mean square error (RMSE) are commonly reported error metrics although there is much debate over which is more appropriate.[200] The MAE provides straightforward information about the average magnitude of errors to be expected from a model, however it does not offer any information into the variability of prediction errors. The RMSE amplifies larger errors due to its squared nature and is sensitive to outliers. Therefore, we considered these metrics together, MAE and RMSE as well as the median absolute error (MdAE) because it is robust to outliers as it is calculated by taking the median of all absolute differences between the target and the prediction.[200] We can also quantify the extent of correlation between the prediction results and ground truth by reporting the coefficient of determination (R^2).

We trained our model on all the information in the train set and evaluated the model on the reserved test set both as a whole and for each individual adsorbate. Reported metrics can be seen pictorially in Figure 21, as well as in Table 5. The MAE across all adsorbates is 2.35 kcal/mol. Of the monatomic adsorbates, C had the smallest MAE at 1.88 kcal/mol. Of the polyatomic adsorbates, the MAEs range from 2.46 to 4.10 kcal/mol. RMSEs are similar although they tend to be 1 to 2 kcal/mol higher, except in the case of *NH, which the RMSE is 5 kcal/mol higher than that of its associated MAE. Table 5 also shows the maximum error on prediction per adsorbate; here it can be seen the *NH has the largest maximum error explaining the difference between the MAE and the RMSE. Figure 22 shows the distribution of prediction errors where it is slightly skewed to the right, explaining the maximum errors. We also choose to report the MdAE which is robust to outliers; the MdAE across all adsorbates is 1.30 kcal/mol, with the largest MdAE for *SH at 2.69 kcal/mol. This alludes to the point made earlier that data preprocessing, in this case cleaning, can have a significant impact on the model. It is clear there are outlier systems in our dataset; the appropriate action to take to filter out this information requires further study. Here, in all reported metrics and figures, we included all data points. Figure 20 shows the prediction values as a function of the test target values to visually show the correlation between the test and prediction for various adsorbates, which are indicated by shape and color. Figure 20 also shows coefficients of determination, where they range from 0.999 to 0.715, all indicating a strong relationship between the prediction and ground truth. As shown the predictions for the monatomic adsorbates are more related to the test than that of the polyatomic adsorbates. This could have occurred for several reasons; there is more data concerning monoatomic adsorption in the database as seen in Table 4, and the targets for the polyatomic adsorption tend to span a smaller range of values which may indicate that this distribution of values may not have been ideal for the chosen model. As evidence, the correlation coefficient for water adsorption stands out as being particularly poor; here the adsorption energies of water span the narrowest distribution as seen by Table 6.

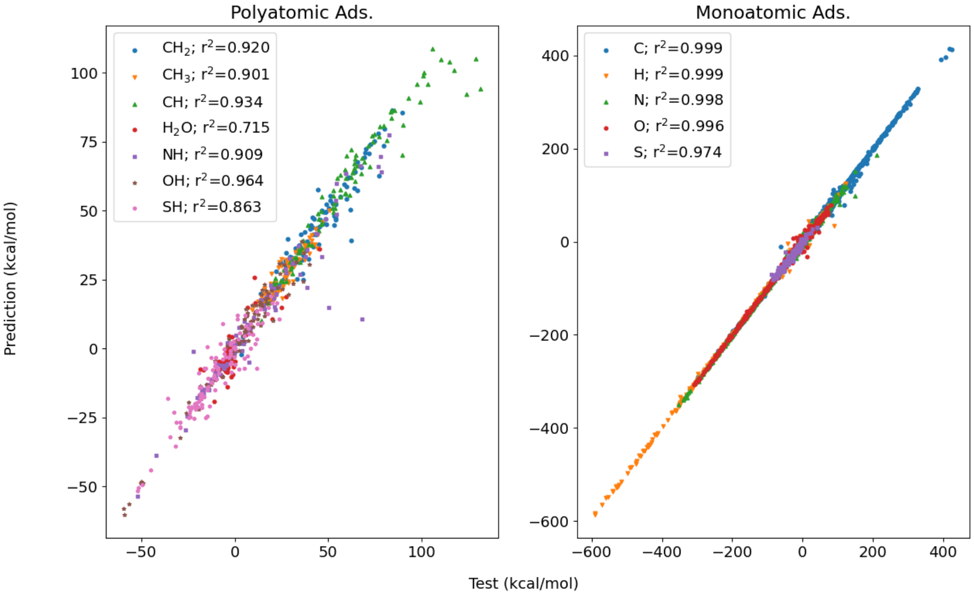


Figure 20 : Test values versus prediction in kcal/mol for various adsorbates split between monatomic and polyatomic adsorbates. Correlation coefficients are shown between the test and prediction.

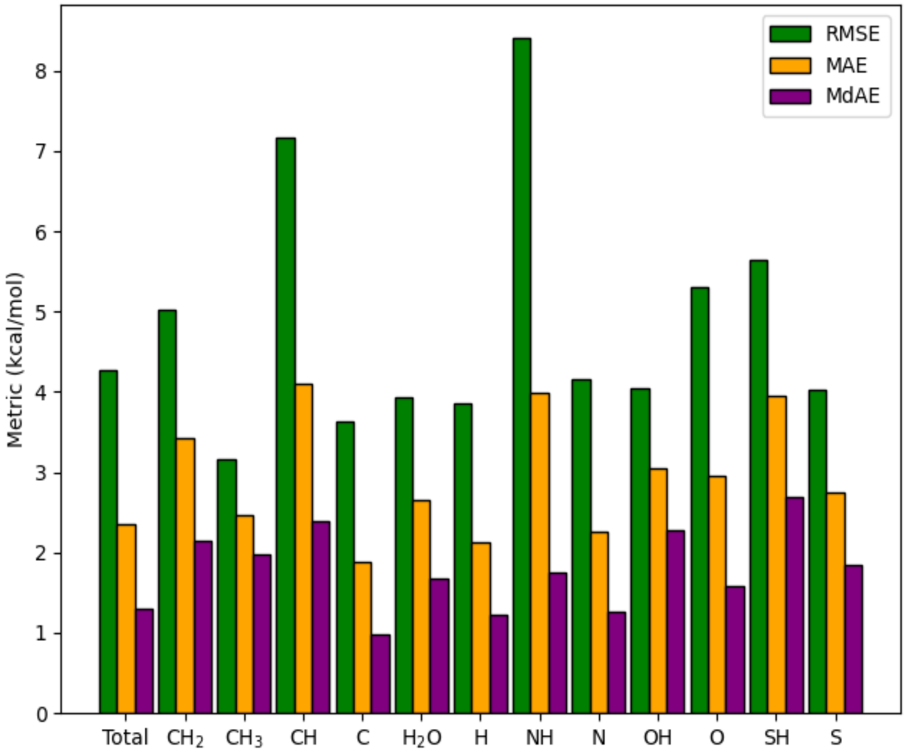


Figure 21 : Metric values as evaluated on the reserved test set, shown as total and split by adsorbate. RMSE is root mean squared error, MAE is mean absolute error, MdAE is median absolute error.

Figure 22 : A histogram of the prediction error (y_{pred} - y_{test}) in kcal/mol.

Table 4 : Number of each adsorbate type found in total dataset

	CH_2	CH_3	CH	C	H_2O	H	NH	N	OH	O	SH	S
Number of Adsorbates	1007	864	1035	17034	1025	12951	1066	7135	761	5226	1223	3465

Table 5 : Metrics as evaluated on the reserved test set

kcal/mol	Overall	CH_2	CH_3	CH	C	H_2O	H	NH	N	OH	O	SH	S
RMSE	4.27	5.02	3.17	7.16	3.63	3.93	3.86	8.40	4.16	4.04	5.31	5.65	4.03
MAE	2.35	3.42	2.46	4.11	1.88	2.65	2.13	3.98	2.26	3.05	2.96	3.94	2.74
MdAE	1.30	2.15	1.97	2.39	0.98	1.67	1.22	1.74	1.25	2.28	1.57	2.69	1.85
MAX ERROR	57.05	23.28	7.60	37.27	49.92	15.05	55.34	57.05	49.51	12.26	45.45	19.00	26.01

Table 6 : Maximum, minimum, and range of targets per adsorbate

kcal/mol	CH_2	CH_3	CH	C	H_2O	H	NH	N	OH	O	SH	S
Maximum	107.7	52.8	146.8	473.1	46.2	316.1	92.1	289.9	48.4	144.5	37.7	43.9
Minimum	-13.4	-24.2	-4.02	-453.3	-21.3	-591.4	-72.9	-362.9	-63.0	-314.5	-70.0	-98.6
Range	121.1	77.1	150.9	926.4	67.5	907.5	165.0	652.8	111.5	459.0	107.7	142.5

11.3.2 Analysis of ML features

After developing a ML model, we wished to understand which of the features considered are most important to determining the adsorption energy. We utilized a unified framework to interpret model predictions employing Shapley Additive Explanations (SHAP),[201] which is a game theoretic approach to explain the output of any machine learning model, in line with other groups in this field.[202] The mean absolute SHAP value for each feature quantifies the magnitude of each feature's contribution towards the target. Features with higher mean absolute SHAP values are more influential, in this way the mean absolute SHAP values can be thought of as a measure of feature importance.[201] Figure 23(a) shows the mean absolute SHAP value for the ten most influential features where seven of these features come from the embedded nominal information, highlighting this data's importance in the model. The Beeswarm plot in Figure 23(b) shows the relationship between the actual feature value and its impact on the model as well as the distribution of values. For example, we see that the value of "ads_H_1" is directly related to its importance in the model. Here, it can be easily seen that the three most important features are the "ads_H_1", "ads_2_embedding_0", and "energy_per_atom" which relate to the enthalpy of the adsorbate, an entity embedded feature, and the energy of the surface per atom, respectively. This is encouraging as it indicates that model developed a relationship between the surface and the adsorbate using the categorical information.

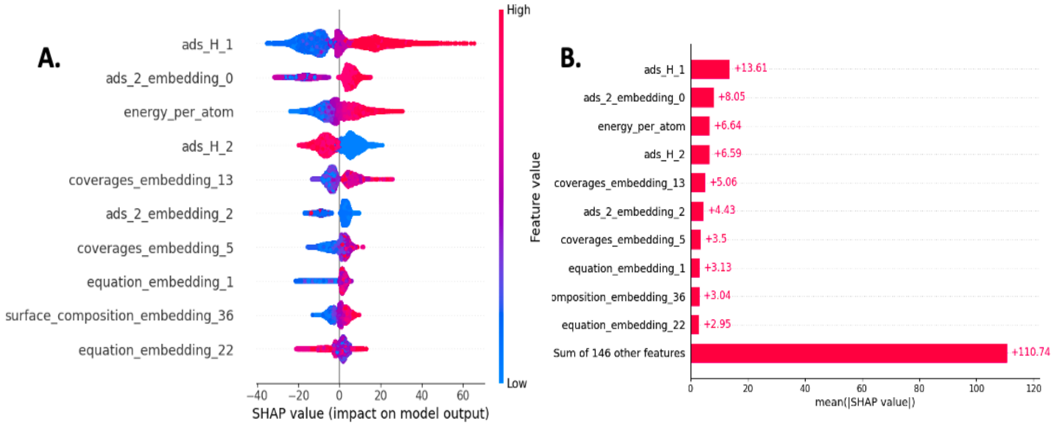


Figure 23 : SHAP values in A. Beeswarm plot with the feature value represented by color B. the mean shape values shown for the top 10 features.

11.4 Conclusion

Machine learning methods in combination with *ab initio* methods have created the new frontier in understanding catalytic reactions and devising new catalysts. Through continuous development in catalysis theory and break throughs in ML techniques, we encourage and expect more data-driven approaches be employed to understand and predict catalytic reactions over surfaces. To conclude, a machine-learning framework is explored and developed to accelerate the discovery of catalysts by learning adsorption energies. This approach is unique in that it uses nominal categorical and readily available information so that it can rapidly explore broad chemical space with minimal user effort. By learning from the data of 52,792 alloys the adsorption energies of twelve diverse adsorbates, can be predicted for new bimetallic catalysts with a MAE of 2.35 kcal/mol. Feature analysis indicates that the most important features used in the building of this model were from the entity embedding process. Given recent machine learning advances in the field of natural language processing, this work highlights the necessity of exploring the use of alternative data sources in chemical space.

12 Conclusion

The investigation of mechanisms concerning energy related materials has seen significant progress as presented here. These advances and recent improvements in data science related methods for chemical purposes have paved the way for discovering new catalysts. In chapter 2, I discussed the investigation of a multifunctional magnetic metal-organic framework whose precedent shows promise in enabling new applications in green technologies and next-generation data processing and storage. In chapters 3 and 4, I focused on elucidating mechanisms concerning the nitrogen reduction reaction and the methane oxidation reaction on surfaces. Both reactions are of exceptional scientific, societal, and industrial importance. These chapters employ methods like that used in the data generation of the information utilized in chapter 5, where I developed a novel way to identify catalytic materials.

Overall, approximations such as linear scaling relations, capitalized on in chapter 5, have played a pivotal role for efficiently designing new catalysts. I anticipate the next generation of research efforts will push towards breaking these scaling relations, thus creating more generalizable models across chemical phenomena. This work would involve the creation of methods that more rigorously reproduce experimental reaction environments. I hope that developing more experimentally relevant models will break linear scaling relationships, especially for well-studied and industrially relevant reactions where exceedingly novel solutions are required to improve existing established catalysts. The community will need to invest in building capabilities for probing atomic-scale phenomena and reaction mechanisms in the presence of complex reaction environments. As an example, a plethora of work remains in the area of solvent modeling, especially with more complicated solvents such as ionic liquids.[203] Further examples include reaction environments which foster competing mechanisms, as discussed in chapter 4. Limitations like these frustrate the atomic-level understanding of these catalysts despite their capacity for use in many industrially significant reactions. For these developments to come to fruition the computational catalysis community will need to work synergistically with the experimental community in studying complex systems.

With time, I envision challenges in using said developed information to its full potential. Despite significant advances in the implementation of ML techniques on catalytic systems, there are still several challenges that need to be addressed. The development of relevant datasets that accurately reflect the intrinsic nature of the catalyst activity, is of paramount importance. This lack of information provides a huge area of growth as it is a guess as to what information is most needed to construct these models. ML is poised to be a versatile quantitative tool to study and understand catalytic processes.

13 Bibliography

(1) Singh, D. J.; Nordström, L. *Planewaves, Pseudopotentials, and the LAPW Method*, 2nd ed.; Springer: New York, NY, 2006.

(2) Quantum Mechanics of Many-Electron Systems. 20.

(3) Slater, J. C. A Simplification of the Hartree-Fock Method. *Phys. Rev.* **1951**, *81* (3), 385–390. https://doi.org/10.1103/PhysRev.81.385.

(4) Hohenberg, P.; Kohn, W. Inhomogeneous Electron Gas. *Phys. Rev.* **1964**, *136* (3B), B864–B871. https://doi.org/10.1103/PhysRev.136.B864.

(5) Zunger, A.; Freeman, A. J. Ground-State Electronic Properties of Diamond in the Local-Density Formalism. *Phys. Rev. B* **1977**, *15* (10), 5049–5065. https://doi.org/10.1103/PhysRevB.15.5049.

(6) Zunger, A.; Freeman, A. J. Self-Consistent Numerical-Basis-Set Linear-Combination-of-Atomic-Orbitals Investigation of the Electronic Structure and Properties of Ti S 2. *Phys. Rev. B* **1977**, *16* (2), 906–924. https://doi.org/10.1103/PhysRevB.16.906.

(7) Zunger, A.; Freeman, A. J. Ground- and Excited-State Properties of LiF in the Local-Density Formalism. *Phys. Rev. B* **1977**, *16* (6), 2901–2926. https://doi.org/10.1103/PhysRevB.16.2901.

(8) Ernzerhof, M.; Perdew, J. P.; Burke, K. Density Functionals: Where Do They Come from, Why Do They Work? In *Density Functional Theory I*; Nalewajski, R. F., Ed.; Dunitz, J. D., Hafner, K., Houk, K. N., Ito, S., Lehn, J.-M.,

Raymond, K. N., Rees, C. W., Thiem, J., Vögtle, F., Series Eds.; Topics in Current Chemistry; Springer Berlin Heidelberg: Berlin, Heidelberg, 1996; Vol. 180, pp 1–30. https://doi.org/10.1007/3-540-61091-X_1.

(9) Kohn, W.; Sham, L. J. Self-Consistent Equations Including Exchange and Correlation Effects. *Phys. Rev.* **1965**, *140* (4A), A1133–A1138. https://doi.org/10.1103/PhysRev.140.A1133.

(10) Gutfleisch, O.; Willard, M. A.; Brück, E.; Chen, C. H.; Sankar, S. G.; Liu, J. P. Magnetic Materials and Devices for the 21st Century: Stronger, Lighter, and More Energy Efficient. *Adv. Mater.* **2011**, *23* (7), 821–842. https://doi.org/10.1002/adma.201002180.

(11) Felser, C.; Fecher, G. H.; Balke, B. Spintronics: A Challenge for Materials Science and Solid-State Chemistry. *Angew. Chem. Int. Ed.* **2007**, *46* (5), 668–699. https://doi.org/10.1002/anie.200601815.

(12) Coronado, E.; Palacio, F.; Veciana, J. Molecule-Based Magnetic Materials. *Angew. Chem. Int. Ed.* **2003**, *42* (23), 2570–2572. https://doi.org/10.1002/anie.200390487.

(13) Kübler, J. *Theory of Itinerant Electron Magnetism*; Oxford Univ. Press, 2000.

(14) Zener, C. Interaction between the d -Shells in the Transition Metals. II. Ferromagnetic Compounds of Manganese with Perovskite Structure. *Phys. Rev.* **1951**, *82* (3), 403–405. https://doi.org/10.1103/PhysRev.82.403.

(15) Briceño, G.; Chang, H.; Sun, X.; Schultz, P. G.; Xiang, X.-D. A Class of Cobalt Oxide Magnetoresistance Materials Discovered with Combinatorial Synthesis. *Science* **1995**, *270* (5234), 273–275. https://doi.org/10.1126/science.270.5234.273.

(16) Şaşıoğlu, E.; Sandratskii, L. M.; Bruno, P. Role of Conduction Electrons in Mediating Exchange Interactions in Mn-Based Heusler Alloys. *Phys. Rev. B* **2008**, *77* (6), 064417. https://doi.org/10.1103/PhysRevB.77.064417.

(17) Yaghi, O. M.; O'Keeffe, M.; Ockwig, N. W.; Chae, H. K.; Eddaoudi, M.; Kim, J. Reticular Synthesis and the Design of New Materials. *Nature* **2003**, *423* (6941), 705–714. https://doi.org/10.1038/nature01650.

(18) Calbo, J.; Golomb, M. J.; Walsh, A. Redox-Active Metal–Organic Frameworks for Energy Conversion and Storage. *J. Mater. Chem. A* **2019**, *7* (28), 16571–16597. https://doi.org/10.1039/C9TA04680A.

(19) Yin, Z.; Wan, S.; Yang, J.; Kurmoo, M.; Zeng, M.-H. Recent Advances in Post-Synthetic Modification of Metal–Organic Frameworks: New Types and Tandem Reactions. *Coord. Chem. Rev.* **2019**, *378*, 500–512. https://doi.org/10.1016/j.ccr.2017.11.015.

(20) Dechambenoit, P.; Long, J. R. Microporous Magnets. *Chem. Soc. Rev.* **2011**, *40* (6), 3249. https://doi.org/10.1039/c0cs00167h.

(21) Thorarinsdottir, A. E.; Harris, T. D. Metal–Organic Framework Magnets. *Chem. Rev.* **2020**, *120* (16), 8716–8789. https://doi.org/10.1021/acs.chemrev.9b00666.

(22) Kosaka, W.; Liu, Z.; Zhang, J.; Sato, Y.; Hori, A.; Matsuda, R.; Kitagawa, S.; Miyasaka, H. Gas-Responsive Porous Magnet Distinguishes the Electron Spin of Molecular Oxygen. *Nat. Commun.* **2018**, *9* (1), 5420. https://doi.org/10.1038/s41467-018-07889-1.

(23) Ruiz, E.; Rodríguez-Fortea, A.; Alvarez, S.; Verdaguer, M. Is It Possible To Get High T_C Magnets with Prussian Blue Analogues? A Theoretical Prospect. *Chem. - Eur. J.* **2005**, *11* (7), 2135–2144. https://doi.org/10.1002/chem.200400821.

(24) Mallah, T.; Thiébaut, S.; Verdaguer, M.; Veillet, P. High- T_c Molecular-Based Magnets: Ferrimagnetic Mixed-Valence Chromium(III)-Chromium(II) Cyanides with T_c at 240 and 190 Kelvin. *Science* **1993**, *262* (5139), 1554–1557. https://doi.org/10.1126/science.262.5139.1554.

(25) Ferlay, S.; Mallah, T.; Ouahès, R.; Veillet, P.; Verdaguer, M. A Room-Temperature Organometallic Magnet Based on Prussian Blue. *Nature* **1995**, *378* (6558), 701–703. https://doi.org/10.1038/378701a0.

(26) Holmes, S. M.; Girolami, G. S. Sol−Gel Synthesis of $KV^{II}[Cr^{III}(CN)_6]\cdot 2H_2O$: A Crystalline Molecule-Based Magnet with a Magnetic Ordering Temperature above 100 °C. *J. Am. Chem. Soc.* **1999**, *121* (23), 5593–5594. https://doi.org/10.1021/ja990946c.

(27) Manriquez, J. M.; Yee, G. T.; McLean, R. S.; Epstein, A. J.; Miller, J. S. A Room-Temperature Molecular/Organic-Based Magnet. *Science* **1991**, *252* (5011), 1415–1417. https://doi.org/10.1126/science.252.5011.1415.

(28) Bechlars, B.; D'Alessandro, D. M.; Jenkins, D. M.; Iavarone, A. T.; Glover, S. D.; Kubiak, C. P.; Long, J. R. High-Spin Ground States via Electron Delocalization in Mixed-Valence Imidazolate-Bridged Divanadium Complexes. *Nat. Chem.* **2010**, *2* (5), 362–368. https://doi.org/10.1038/nchem.585.

(29) Gaudette, A. I.; Jeon, I.-R.; Anderson, J. S.; Grandjean, F.; Long, G. J.; Harris, T. D. Electron Hopping through Double-Exchange Coupling in a Mixed-Valence Diiminobenzoquinone-Bridged Fe $_2$ Complex. *J. Am. Chem. Soc.* **2015**, *137* (39), 12617–12626. https://doi.org/10.1021/jacs.5b07251.

(30) Schulze, B.; Schubert, U. S. Beyond Click Chemistry – Supramolecular Interactions of 1,2,3-Triazoles. *Chem. Soc. Rev.* **2014**, *43* (8), 2522. https://doi.org/10.1039/c3cs60386e.

(31) Aubrey, M. L.; Wiers, B. M.; Andrews, S. C.; Sakurai, T.; Reyes-Lillo, S. E.; Hamed, S. M.; Yu, C.-J.; Darago, L. E.; Mason, J. A.; Baeg, J.-O.; Grandjean, F.; Long, G. J.; Seki, S.; Neaton, J. B.; Yang, P.; Long, J. R. Electron Delocalization and Charge Mobility as a Function of Reduction in a Metal–Organic Framework. *Nat. Mater.* **2018**, *17* (7), 625–632. https://doi.org/10.1038/s41563-018-0098-1.

(32) Zhang, J.-P.; Zhang, Y.-B.; Lin, J.-B.; Chen, X.-M. Metal Azolate Frameworks: From Crystal Engineering to Functional Materials. *Chem. Rev.* **2012**, *112* (2), 1001–1033. https://doi.org/10.1021/cr200139g.

(33) Park, J. G.; Aubrey, M. L.; Oktawiec, J.; Chakarawet, K.; Darago, L. E.; Grandjean, F.; Long, G. J.; Long, J. R. Charge Delocalization and Bulk Electronic

Conductivity in the Mixed-Valence Metal–Organic Framework Fe(1,2,3-Triazolate)$_2$(BF$_4$)$_x$. *J. Am. Chem. Soc.* **2018**, *140* (27), 8526–8534. https://doi.org/10.1021/jacs.8b03696.

(34) Brunschwig, B. S.; Creutz, C.; Sutin, N. Optical Transitions of Symmetrical Mixed-Valence Systems in the Class II–III Transition RegimeElectronic Supplementary Information (ESI) Is Available: Derivation of Eqn. (39c), Table Summarizing the Relationships between Band Maxima and Band Widths Predicted by the Two-State Model and Table of Spectral Properties of Mixed-Valence Ruthenium(II)/(III) Bridged by Pyrazine and Dicyanamide. See Http://Www.Rsc.Org/Suppdata/Cs/B0/B008034i/. *Chem. Soc. Rev.* **2002**, *31* (3), 168–184. https://doi.org/10.1039/b008034i.

(35) Kresse, G.; Furthmüller, J. Efficiency of Ab-Initio Total Energy Calculations for Metals and Semiconductors Using a Plane-Wave Basis Set. *Comput. Mater. Sci.* **1996**, *6* (1), 15–50. https://doi.org/10.1016/0927-0256(96)00008-0.

(36) Kresse, G.; Joubert, D. From Ultrasoft Pseudopotentials to the Projector Augmented-Wave Method. *Phys. Rev. B* **1999**, *59* (3), 1758–1775. https://doi.org/10.1103/PhysRevB.59.1758.

(37) Perdew, J. P.; Burke, K.; Ernzerhof, M. Generalized Gradient Approximation Made Simple. *Phys. Rev. Lett.* **1996**, *77* (18), 3865–3868. https://doi.org/10.1103/PhysRevLett.77.3865.

(38) Manz, T. A.; Sholl, D. S. Methods for Computing Accurate Atomic Spin Moments for Collinear and Noncollinear Magnetism in Periodic and Nonperiodic Materials. *J. Chem. Theory Comput.* **2011**, *7* (12), 4146–4164. https://doi.org/10.1021/ct200539n.

(39) Manz, T. A.; Sholl, D. S. Improved Atoms-in-Molecule Charge Partitioning Functional for Simultaneously Reproducing the Electrostatic Potential and Chemical States in Periodic and Nonperiodic Materials. *J. Chem. Theory Comput.* **2012**, *8* (8), 2844–2867. https://doi.org/10.1021/ct3002199.

(40) Manz, T. A.; Sholl, D. S. Chemically Meaningful Atomic Charges That Reproduce the Electrostatic Potential in Periodic and Nonperiodic Materials. *J. Chem. Theory Comput.* **2010**, *6* (8), 2455–2468. https://doi.org/10.1021/ct100125x.

(41) Jmol: An Open-Source Java Viewer for Chemical Structures in 3D. http://www.jmol.org/.

(42) Momma, K.; Izumi, F. *VESTA* : A Three-Dimensional Visualization System for Electronic and Structural Analysis. *J. Appl. Crystallogr.* **2008**, *41* (3), 653–658. https://doi.org/10.1107/S0021889808012016.

(43) Murray, L. J.; Dinca, M.; Yano, J.; Chavan, S.; Bordiga, S.; Brown, C. M.; Long, J. R. Highly-Selective and Reversible O_2 Binding in Cr_3 (1,3,5-Benzenetricarboxylate)$_2$. *J. Am. Chem. Soc.* **2010**, *132* (23), 7856–7857. https://doi.org/10.1021/ja1027925.

(44) Kulik, H. J. Perspective: Treating Electron over-Delocalization with the DFT+U Method. *J. Chem. Phys.* **2015**, *142* (24), 240901. https://doi.org/10.1063/1.4922693.

(45) Ong, S. P.; Richards, W. D.; Jain, A.; Hautier, G.; Kocher, M.; Cholia, S.; Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Genomics (Pymatgen): A Robust, Open-Source Python Library for Materials Analysis. *Comput. Mater. Sci.* **2013**, *68*, 314–319. https://doi.org/10.1016/j.commatsci.2012.10.028.

(46) Hunter, J. D. Matplotlib: A 2D Graphics Environment. *Comput. Sci. Eng.* **2007**, *9* (3), 90–95. https://doi.org/10.1109/MCSE.2007.55.

(47) M Ganose, A.; J Jackson, A.; O Scanlon, D. Sumo: Command-Line Tools for Plotting and Analysis of Periodic *Ab Initio* Calculations. *J. Open Source Softw.* **2018**, *3* (28), 717. https://doi.org/10.21105/joss.00717.

(48) Psofogiannakis, G.; St-Amant, A.; Ternan, M. Methane Oxidation Mechanism on Pt(111): A Cluster Model DFT Study. *J. Phys. Chem. B* **2006**, *110* (48), 24593–24605. https://doi.org/10.1021/jp061559+.

(49) Duinen, G. V. How a Century of Ammonia Synthesis Changed the World. *Nat. Geosci.* **2008**, *1*, 4.

(50) Rosca, V.; Duca, M.; de Groot, M. T.; Koper, M. T. M. Nitrogen Cycle Electrocatalysis. *Chem. Rev.* **2009**, *109* (6), 2209–2244. https://doi.org/10.1021/cr8003696.

(51) Green, D. W. *Perry's Chemical Engineers' Handbook*; McGraw-Hill: New York, NY, 2008; Vol. 796.

(52) Kuang, X.-J.; Wang, X.-Q.; Liu, G.-B. A Density Functional Study on the Adsorption of Hydrogen Molecule onto Small Copper Clusters. *J. Chem. Sci.* **2011**, *123* (5), 743–754. https://doi.org/10.1007/s12039-011-0130-3.

(53) Soon, A.; Wong, L.; Lee, M.; Todorova, M.; Delley, B.; Stampfl, C. Nitrogen Adsorption and Thin Surface Nitrides on Cu(111) from First-Principles. *Surf. Sci.* **2007**, *601* (21), 4775–4785. https://doi.org/10.1016/j.susc.2007.07.011.

(54) Sangiovanni, D. G.; Mei, A. B.; Hultman, L.; Chirita, V.; Petrov, I.; Greene, J. E. *Ab Initio* Molecular Dynamics Simulations of Nitrogen/VN(001) Surface Reactions: Vacancy-Catalyzed N_2 Dissociative Chemisorption, N Adatom Migration, and N_2 Desorption. *J. Phys. Chem. C* **2016**, *120* (23), 12503–12516. https://doi.org/10.1021/acs.jpcc.6b02652.

(55) Abghoui, Y.; Skúlason, E. Hydrogen Evolution Reaction Catalyzed by Transition-Metal Nitrides. *J. Phys. Chem. C* **2017**, *121* (43), 24036–24045. https://doi.org/10.1021/acs.jpcc.7b06811.

(56) Höskuldsson, Á. B.; Abghoui, Y.; Gunnarsdóttir, A. B.; Skúlason, E. Computational Screening of Rutile Oxides for Electrochemical Ammonia Formation. *ACS Sustain. Chem. Eng.* **2017**, *5* (11), 10327–10333. https://doi.org/10.1021/acssuschemeng.7b02379.

(57) Abghoui, Y.; Skúlason, E. Computational Predictions of Catalytic Activity of Zincblende (110) Surfaces of Metal Nitrides for Electrochemical Ammonia

Synthesis. *J. Phys. Chem. C* **2017**, *121* (11), 6141–6151. https://doi.org/10.1021/acs.jpcc.7b00196.

(58) Abghoui, Y.; Skúlason, E. Electrochemical Synthesis of Ammonia via Mars-van Krevelen Mechanism on the (111) Facets of Group III–VII Transition Metal Mononitrides. *Catal. Today* **2017**, *286*, 78–84. https://doi.org/10.1016/j.cattod.2016.06.009.

(59) Abghoui, Y.; Garden, A. L.; Howalt, J. G.; Vegge, T.; Skúlason, E. Electroreduction of N_2 to Ammonia at Ambient Conditions on Mononitrides of Zr, Nb, Cr, and V: A DFT Guide for Experiments. *ACS Catal.* **2016**, *6* (2), 635–646. https://doi.org/10.1021/acscatal.5b01918.

(60) Abghoui, Y.; Skúlason, E. Onset Potentials for Different Reaction Mechanisms of Nitrogen Activation to Ammonia on Transition Metal Nitride Electro-Catalysts. *Catal. Today* **2017**, *286*, 69–77. https://doi.org/10.1016/j.cattod.2016.11.047.

(61) Rostamikia, G.; Maheshwari, S.; Janik, M. J. Elementary Kinetics of Nitrogen Electroreduction to Ammonia on Late Transition Metals. *Catal. Sci. Technol.* **2019**, *9* (1), 174–181. https://doi.org/10.1039/C8CY01845F.

(62) Skúlason, E.; Bligaard, T.; Gudmundsdóttir, S.; Studt, F.; Rossmeisl, J.; Abild-Pedersen, F.; Vegge, T.; Jónsson, H.; Nørskov, J. K. A Theoretical Evaluation of Possible Transition Metal Electro-Catalysts for N_2 Reduction. *Phys Chem Chem Phys* **2012**, *14* (3), 1235–1245. https://doi.org/10.1039/C1CP22271F.

(63) Montoya, J. H.; Persson, K. A. A High-Throughput Framework for Determining Adsorption Energies on Solid Surfaces. *Npj Comput. Mater.* **2017**, *3* (1), 14. https://doi.org/10.1038/s41524-017-0017-z.

(64) Singh, M. R.; Goodpaster, J. D.; Weber, A. Z.; Head-Gordon, M.; Bell, A. T. Mechanistic Insights into Electrochemical Reduction of CO $_2$ over Ag Using Density Functional Theory and Transport Models. *Proc. Natl. Acad. Sci.* **2017**, *114* (42). https://doi.org/10.1073/pnas.1713164114.

(65) Zijlstra, B.; Zhang, X.; Liu, J.-X.; Filot, I. A. W.; Zhou, Z.; Sun, S.; Hensen, E. J. M. First-Principles Microkinetics Simulations of Electrochemical Reduction of CO2 over Cu Catalysts. *Electrochimica Acta* **2020**, *335*, 135665. https://doi.org/10.1016/j.electacta.2020.135665.

(66) Skúlason, E.; Tripkovic, V.; Björketun, M. E.; Gudmundsdóttir, S.; Karlberg, G.; Rossmeisl, J.; Bligaard, T.; Jónsson, H.; Nørskov, J. K. Modeling the Electrochemical Hydrogen Oxidation and Evolution Reactions on the Basis of Density Functional Theory Calculations. *J. Phys. Chem. C* **2010**, *114* (42), 18182–18197. https://doi.org/10.1021/jp1048887.

(67) Bagherzadeh Mostaghimi, A. H.; Al-Attas, T. A.; Kibria, M. G.; Siahrostami, S. A Review on Electrocatalytic Oxidation of Methane to Oxygenates. *J. Mater. Chem. A* **2020**, *8* (31), 15575–15590. https://doi.org/10.1039/D0TA03758C.

(68) Guene Lougou, B.; Shuai, Y.; Chaffa, G.; Ahouannou, C.; Pan, R.; Zhang, H.; Tan, H. Analysis of Two-Step Solar Thermochemical Looping Reforming of

Fe_3O_4 Redox Cycles for Synthesis Gas Production. *Energy Technol.* **2019**, *7* (3), 1800588. https://doi.org/10.1002/ente.201800588.

(69) Jang, J.; Shen, K.; Morales-Guio, C. G. Electrochemical Direct Partial Oxidation of Methane to Methanol. *Joule* **2019**, *3* (11), 2589–2593. https://doi.org/10.1016/j.joule.2019.10.004.

(70) Olah, G. A. Beyond Oil and Gas: The Methanol Economy. *Angew. Chem. Int. Ed.* **2005**, *44* (18), 2636–2639. https://doi.org/10.1002/anie.200462121.

(71) Ravi, M.; Ranocchiari, M.; van Bokhoven, J. A. The Direct Catalytic Oxidation of Methane to Methanol-A Critical Assessment. *Angew. Chem. Int. Ed.* **2017**, *56* (52), 16464–16483. https://doi.org/10.1002/anie.201702550.

(72) Go, D.; Yang, B. C.; Shin, J. W.; Kim, H. J.; Lee, S.; Kye, S.; Kim, S.; An, J. Atomic Layer Deposited YSZ Overlayer on Ru for Direct Methane Utilization in Solid Oxide Fuel Cell. *Ceram. Int.* **2020**, *46* (2), 1705–1710. https://doi.org/10.1016/j.ceramint.2019.09.143.

(73) Prakash, G. K. S.; Smart, M. C.; Wang, Q.-J.; Atti, A.; Pleynet, V.; Yang, B.; McGrath, K.; Olah, G. A.; Narayanan, S. R.; Chun, W.; Valdez, T.; Surampudi, S. High Efficiency Direct Methanol Fuel Cell Based on Poly(Styrenesulfonic) Acid (PSSA)–Poly(Vinylidene Fluoride) (PVDF) Composite Membranes. *J. Fluor. Chem.* **2004**, *125* (8), 1217–1230. https://doi.org/10.1016/j.jfluchem.2004.05.019.

(74) Schwach, P.; Pan, X.; Bao, X. Direct Conversion of Methane to Value-Added Chemicals over Heterogeneous Catalysts: Challenges and Prospects. *Chem. Rev.* **2017**, *117* (13), 8497–8520. https://doi.org/10.1021/acs.chemrev.6b00715.

(75) Thakkar, A. J.; Wu, T. How Well Do Static Electronic Dipole Polarizabilities from Gas-Phase Experiments Compare with Density Functional and MP2 Computations? *J. Chem. Phys.* **2015**, *143* (14), 144302. https://doi.org/10.1063/1.4932594.

(76) Yamamoto, S.; Alcauskas, J. B.; Crozier, T. E. Solubility of Methane in Distilled Water and Seawater. *J. Chem. Eng. Data* **1976**, *21* (1), 78–80. https://doi.org/10.1021/je60068a029.

(77) Baltrusaitis, J.; Jansen, I.; Schuttlefield Christus, J. D. Renewable Energy Based Catalytic CH4 Conversion to Fuels. *Catal. Sci. Technol.* **2014**, *4* (8), 2397. https://doi.org/10.1039/c4cy00294f.

(78) Spinner, N.; Mustain, W. E. Electrochemical Methane Activation and Conversion to Oxygenates at Room Temperature. *ECS Trans.* **2013**, *53* (23), 1–20. https://doi.org/10.1149/05323.0001ecst.

(79) Fornaciari, J. C.; Primc, D.; Kawashima, K.; Wygant, B. R.; Verma, S.; Spanu, L.; Mullins, C. B.; Bell, A. T.; Weber, A. Z. A Perspective on the Electrochemical Oxidation of Methane to Methanol in Membrane Electrode Assemblies. *ACS Energy Lett.* **2020**, *5* (9), 2954–2963. https://doi.org/10.1021/acsenergylett.0c01508.

(80) Murray, E. P.; Tsai, T.; Barnett, S. A. A Direct-Methane Fuel Cell with a Ceria-Based Anode. *Nature* **1999**, *400* (6745), 649–651. https://doi.org/10.1038/23220.

(81) Wang, J. B.; Jang, J.-C.; Huang, T.-J. Study of Ni-Samaria-Doped Ceria Anode for Direct Oxidation of Methane in Solid Oxide Fuel Cells. *J. Power Sources* **2003**, *122* (2), 122–131. https://doi.org/10.1016/S0378-7753(03)00438-5.

(82) Zhan, Z.; Lin, Y.; Pillai, M.; Kim, I.; Barnett, S. A. High-Rate Electrochemical Partial Oxidation of Methane in Solid Oxide Fuel Cells. *J. Power Sources* **2006**, *161* (1), 460–465. https://doi.org/10.1016/j.jpowsour.2006.04.139.

(83) Liu, J. Operation of Anode-Supported Solid Oxide Fuel Cells on Methane and Natural Gas. *Solid State Ion.* **2003**, *158* (1–2), 11–16. https://doi.org/10.1016/S0167-2738(02)00769-5.

(84) Liu, C.; Ye, J.; Jiang, J.; Pan, Y. Progresses in the Preparation of Coke Resistant Ni-Based Catalyst for Steam and CO2 Reforming of Methane. *ChemCatChem* **2011**, *3* (3), 529–541. https://doi.org/10.1002/cctc.201000358.

(85) Ideris, A.; Croiset, E.; Pritzker, M.; Amin, A. Direct-Methane Solid Oxide Fuel Cell (SOFC) with Ni-SDC Anode-Supported Cell. *Int. J. Hydrog. Energy* **2017**, *42* (36), 23118–23129. https://doi.org/10.1016/j.ijhydene.2017.07.117.

(86) Putna, E. S.; Stubenrauch, J.; Vohs, J. M.; Gorte, R. J. Ceria-Based Anodes for the Direct Oxidation of Methane in Solid Oxide Fuel Cells. *Langmuir* **1995**, *11* (12), 4832–4837. https://doi.org/10.1021/la00012a040.

(87) Steele, B. Oxidation of Methane in Solid State Electrochemical Reactors. *Solid State Ion.* **1988**, *28–30*, 1547–1552. https://doi.org/10.1016/0167-2738(88)90417-1.

(88) Kim, H.; Lu, C.; Worrell, W. L.; Vohs, J. M.; Gorte, R. J. Cu-Ni Cermet Anodes for Direct Oxidation of Methane in Solid-Oxide Fuel Cells. *J. Electrochem. Soc.* **2002**, *149* (3), A247. https://doi.org/10.1149/1.1445170.

(89) Ding, H.; Tao, Z.; Liu, S.; Yang, Y. A Redox-Stable Direct-Methane Solid Oxide Fuel Cell (SOFC) with Sr2FeNb0.2Mo0.8O6−δ Double Perovskite as Anode Material. *J. Power Sources* **2016**, *327*, 573–579. https://doi.org/10.1016/j.jpowsour.2016.07.101.

(90) Cook, R. L.; Sammells, A. F. Ambient Temperature Methane Activation to Condensed Species under Cathodic Conditions. *J. Electrochem. Soc.* **1990**, *137* (6), 2007–2008. https://doi.org/10.1149/1.2086854.

(91) Otagawa, T.; Zaromb, S.; Stetter, J. R. Electrochemical Oxidation of Methane in Nonaqueous Electrolytes at Room Temperature: Application to Gas Sensors. *J. Electrochem. Soc.* **1985**, *132* (12), 2951–2957. https://doi.org/10.1149/1.2113699.

(92) Stoukides, M. Electrochemical Studies of Methane Activation: Reviews in Applied Electrochemistry No 42. *J. Appl. Electrochem.* **1995**, *25* (10). https://doi.org/10.1007/BF00241584.

(93) Cassidy, J.; Khoo, S. B.; Pons, S.; Fleischmann, M. Electrochemistry at Very High Potentials: The Use of Ultramicroelectrodes in the Anodic Oxidation of Short-Chain Alkanes. *J. Phys. Chem.* **1985**, *89* (18), 3933–3935. https://doi.org/10.1021/j100264a034.

(94) Frese, K. W. Partial Electrochemical Oxidation of Methane under Mild Conditions. *Langmuir* **1991**, *7* (1), 13–15. https://doi.org/10.1021/la00049a004.

(95) Hahn, F.; Melendres, C. A. Anodic Oxidation of Methane at Noble Metal Electrodes: An ‘in Situ’ Surface Enhanced Infrared Spectroelectrochemical Study. *Electrochimica Acta* **2001**, *46* (23), 3525–3534. https://doi.org/10.1016/S0013-4686(01)00649-1.

(96) Ogura, K.; Migita, C. T.; Ito, Y. Combined Photochemical and Electrochemical Oxidation of Methane. *J. Electrochem. Soc.* **1990**, *137* (2), 500–503. https://doi.org/10.1149/1.2086470.

(97) Wang, Z.; Zeng, X. Bis(Trifluoromethylsulfonyl)Imide (NTf_2)-Based Ionic Liquids for Facile Methane Electro-Oxidation on Pt. *J. Electrochem. Soc.* **2013**, *160* (9), H604–H611. https://doi.org/10.1149/2.039309jes.

(98) Nandenha, J.; Fontes, E. H.; Piasentin, R. M.; Fonseca, F. C.; Neto, A. O. Direct Oxidation of Methane at Low Temperature Using Pt/C, Pd/C, Pt/C-ATO and Pd/C-ATO Electrocatalysts Prepared by Sodium Borohydride Reduction Process. *J. Fuel Chem. Technol.* **2018**, *46* (9), 1137–1145. https://doi.org/10.1016/S1872-5813(18)30046-X.

(99) Ma, H.-B.; Sheng, T.; Yu, W.-S.; Ye, J.-Y.; Wan, L.-Y.; Tian, N.; Sun, S.-G.; Zhou, Z.-Y. High Catalytic Activity of Pt(100) for CH_4 Electrochemical Conversion. *ACS Catal.* **2019**, *9* (11), 10159–10165. https://doi.org/10.1021/acscatal.9b02738.

(100) Kim, R. S.; Surendranath, Y. Electrochemical Reoxidation Enables Continuous Methane-to-Methanol Catalysis with Aqueous Pt Salts. *ACS Cent. Sci.* **2019**, *5* (7), 1179–1186. https://doi.org/10.1021/acscentsci.9b00273.

(101) Fan, Q. Non-Faradaic Electrochemical Promotion of Catalytic Methane Reforming for Methanol Production.

(102) Ma, M.; Oh, C.; Kim, J.; Moon, J. H.; Park, J. H. Electrochemical CH4 Oxidation into Acids and Ketones on ZrO2:NiCo2O4 Quasi-Solid Solution Nanowire Catalyst. *Appl. Catal. B Environ.* **2019**, *259*, 118095. https://doi.org/10.1016/j.apcatb.2019.118095.

(103) Rocha, R. S.; Camargo, L. M.; Lanza, M. R. V.; Bertazzoli, R. A Feasibility Study of the Electro-Recycling of Greenhouse Gases: Design and Characterization of a (TiO2/RuO2)/PTFE Gas Diffusion Electrode for the Electrosynthesis of Methanol from Methane. *Electrocatalysis* **2010**, *1* (4), 224–229. https://doi.org/10.1007/s12678-010-0029-7.

(104) O'Reilly, M. E.; Kim, R. S.; Oh, S.; Surendranath, Y. Catalytic Methane Monofunctionalization by an Electrogenerated High-Valent Pd Intermediate. *ACS Cent. Sci.* **2017**, *3* (11), 1174–1179. https://doi.org/10.1021/acscentsci.7b00342.

(105) Szécsényi, Á.; Li, G.; Gascon, J.; Pidko, E. A. Mechanistic Complexity of Methane Oxidation with H_2O_2 by Single-Site Fe/ZSM-5 Catalyst. *ACS Catal.* **2018**, *8* (9), 7961–7972. https://doi.org/10.1021/acscatal.8b01672.

(106) Liao, M.-S.; Zhang, Q.-E. Dissociation of Methane on Different Transition Metals. *J. Mol. Catal. Chem.* **1998**, *136* (2), 185–194. https://doi.org/10.1016/S1381-1169(98)00050-8.

(107) Chen, Y.; Vlachos, D. G. Density Functional Theory Study of Methane Oxidation and Reforming on Pt(111) and Pt(211). *Ind. Eng. Chem. Res.* **2012**, 120918084645004. https://doi.org/10.1021/ie301792g.

(108) Schröder, D.; Schwarz, H. FeO Activates Methane. *Angew. Chem. Int. Ed. Engl.* **1990**, *29* (12), 1433–1434. https://doi.org/10.1002/anie.199014331.

(109) Dong, A.; Xu, G.; Dai, Z.; Yu, A.; Qiu, S.; Sun, C. Single PdO Loaded on Boron Nanosheet for Methane Oxidation: A DFT Study. *Prog. Nat. Sci. Mater. Int.* **2019**, *29* (3), 367–369. https://doi.org/10.1016/j.pnsc.2019.05.005.

(110) Zhao, Z.-J.; Kulkarni, A.; Vilella, L.; Nørskov, J. K.; Studt, F. Theoretical Insights into the Selective Oxidation of Methane to Methanol in Copper-Exchanged Mordenite. *ACS Catal.* **2016**, *6* (6), 3760–3766. https://doi.org/10.1021/acscatal.6b00440.

(111) Fu, G.; Chen, Z.-N.; Xu, X.; Wan, H.-L. Understanding the Reactivity of the Tetrahedrally Coordinated High-Valence d^0 Transition Metal Oxides toward the C−H Bond Activation of Alkanes: A Cluster Model Study. *J. Phys. Chem. A* **2008**, *112* (4), 717–721. https://doi.org/10.1021/jp709651n.

(112) Latimer, A. A.; Kulkarni, A. R.; Aljama, H.; Montoya, J. H.; Yoo, J. S.; Tsai, C.; Abild-Pedersen, F.; Studt, F.; Nørskov, J. K. Understanding Trends in C–

H Bond Activation in Heterogeneous Catalysis. *Nat. Mater.* **2017**, *16* (2), 225–229. https://doi.org/10.1038/nmat4760.

(113) Abild-Pedersen, F.; Greeley, J.; Studt, F.; Rossmeisl, J.; Munter, T. R.; Moses, P. G.; Skúlason, E.; Bligaard, T.; Nørskov, J. K. Scaling Properties of Adsorption Energies for Hydrogen-Containing Molecules on Transition-Metal Surfaces. *Phys. Rev. Lett.* **2007**, *99* (1), 016105. https://doi.org/10.1103/PhysRevLett.99.016105.

(114) Arnarson, L.; Schmidt, P. S.; Pandey, M.; Bagger, A.; Thygesen, K. S.; Stephens, I. E. L.; Rossmeisl, J. Fundamental Limitation of Electrocatalytic Methane Conversion to Methanol. *Phys. Chem. Chem. Phys.* **2018**, *20* (16), 11152–11159. https://doi.org/10.1039/C8CP01476K.

(115) Danilov, F. I.; Velichenko, A. B.; Nishcheryakova, L. N. Electrocatalytic Processes on Pb/PbO2 Electrodes at High Anodic Potential. *Electrochimica Acta* **1994**, *39*, 1603–1605.

(116) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. A Consistent and Accurate *Ab Initio* Parametrization of Density Functional Dispersion Correction (DFT-D) for the 94 Elements H-Pu. *J. Chem. Phys.* **2010**, *132* (15), 154104. https://doi.org/10.1063/1.3382344.

(117) Nørskov, J. K.; Rossmeisl, J.; Logadottir, A.; Lindqvist, L.; Kitchin, J. R.; Bligaard, T.; Jónsson, H. Origin of the Overpotential for Oxygen Reduction at a Fuel-Cell Cathode. *J. Phys. Chem. B* **2004**, *108* (46), 17886–17892. https://doi.org/10.1021/jp047349j.

(118) Mathew, K.; Sundararaman, R.; Letchworth-Weaver, K.; Arias, T. A.; Hennig, R. G. Implicit Solvation Model for Density-Functional Study of Nanocrystal Surfaces and Reaction Pathways. *J. Chem. Phys.* **2014**, *140* (8), 084106. https://doi.org/10.1063/1.4865107.

(119) Lin, Y.-Z.; Sun, J.; Yi, J.; Lin, J.-D.; Chen, H.-B.; Liao, D.-W. Energetics of Chemisorption and Conversion of Methane on Transition Metal Surfaces. *J. Mol. Struct. THEOCHEM* **2002**, *587* (1–3), 63–71. https://doi.org/10.1016/S0166-1280(02)00097-0.

(120) Kresse, G. *Ab Initio* Molecular Dynamics for Liquid Metals. **1995**, 8.

(121) Ping, Y.; Nielsen, R. J.; Goddard, W. A. The Reaction Mechanism with Free Energy Barriers at Constant Potentials for the Oxygen Evolution Reaction at the IrO_2 (110) Surface. *J. Am. Chem. Soc.* **2017**, *139* (1), 149–155. https://doi.org/10.1021/jacs.6b07557.

(122) Farha, O. K.; Eryazici, I.; Jeong, N. C.; Hauser, B. G.; Wilmer, C. E.; Sarjeant, A. A.; Snurr, R. Q.; Nguyen, S. T.; Yazaydın, A. Ö.; Hupp, J. T. Metal–Organic Framework Materials with Ultrahigh Surface Areas: Is the Sky the Limit? *J. Am. Chem. Soc.* **2012**, *134* (36), 15016–15021. https://doi.org/10.1021/ja3055639.

(123) Liu, J.; Chen, L.; Cui, H.; Zhang, J.; Zhang, L.; Su, C.-Y. Applications of Metal–Organic Frameworks in Heterogeneous Supramolecular Catalysis. *Chem Soc Rev* **2014**, *43* (16), 6011–6061. https://doi.org/10.1039/C4CS00094C.

(124) Chughtai, A. H.; Ahmad, N.; Younus, H. A.; Laypkov, A.; Verpoort, F. Metal–Organic Frameworks: Versatile Heterogeneous Catalysts for Efficient Catalytic Organic Transformations. *Chem. Soc. Rev.* **2015**, *44* (19), 6804–6849. https://doi.org/10.1039/C4CS00395K.

(125) Bernales, V.; Ortuño, M. A.; Truhlar, D. G.; Cramer, C. J.; Gagliardi, L. Computational Design of Functionalized Metal–Organic Framework Nodes for Catalysis. *ACS Cent. Sci.* **2018**, *4* (1), 5–19. https://doi.org/10.1021/acscentsci.7b00500.

(126) Wu, X.-P.; Gagliardi, L.; Truhlar, D. G. Cerium Metal–Organic Framework for Photocatalysis. *J. Am. Chem. Soc.* **2018**, *140* (25), 7904–7912. https://doi.org/10.1021/jacs.8b03613.

(127) Wu, X.; Choudhuri, I.; Truhlar, D. G. Computational Studies of Photocatalysis with Metal–Organic Frameworks. *ENERGY Environ. Mater.* **2019**, *2* (4), 251–263. https://doi.org/10.1002/eem2.12051.

(128) Choudhuri, I.; Truhlar, D. G. Photogenerated Charge Separation in a CdSe Nanocluster Encapsulated in a Metal–Organic Framework for Improved Photocatalysis. *J. Phys. Chem. C* **2020**, *124* (16), 8504–8513. https://doi.org/10.1021/acs.jpcc.0c00007.

(129) Li, L.; He, J.; Wang, Y.; Lv, X.; Gu, X.; Dai, P.; Liu, D.; Zhao, X. Metal–Organic Frameworks: A Promising Platform for Constructing Non-Noble Electrocatalysts for the Oxygen-Reduction Reaction. *J. Mater. Chem. A* **2019**, *7* (5), 1964–1988. https://doi.org/10.1039/C8TA11704G.

(130) Wang, X.; Xiao, B.; Li, Y.; Tang, Y.; Liu, F.; Chen, J.; Liu, Y. First-Principles Based Machine Learning Study of Oxygen Evolution Reactions of Perovskite Oxides Using a Surface Center-Environment Feature Model. *Appl. Surf. Sci.* **2020**, *531*, 147323. https://doi.org/10.1016/j.apsusc.2020.147323.

(131) Shah, M. S.; Tsapatsis, M.; Siepmann, J. I. Hydrogen Sulfide Capture: From Absorption in Polar Liquids to Oxide, Zeolite, and Metal–Organic Framework Adsorbents and Membranes. *Chem. Rev.* **2017**, *117* (14), 9755–9803. https://doi.org/10.1021/acs.chemrev.7b00095.

(132) Bloch, E. D.; Queen, W. L.; Krishna, R.; Zadrozny, J. M.; Brown, C. M.; Long, J. R. Hydrocarbon Separations in a Metal-Organic Framework with Open Iron(II) Coordination Sites. *Science* **2012**, *335* (6076), 1606–1610. https://doi.org/10.1126/science.1217544.

(133) Li, H.; Wang, K.; Sun, Y.; Lollar, C. T.; Li, J.; Zhou, H.-C. Recent Advances in Gas Storage and Separation Using Metal–Organic Frameworks. *Mater. Today* **2018**, *21* (2), 108–121. https://doi.org/10.1016/j.mattod.2017.07.006.

(134) Li, H.; Eddaoudi, M.; O'Keeffe, M.; Yaghi, O. M. Design and Synthesis of an Exceptionally Stable and Highly Porous Metal-Organic Framework. *Nature* **1999**, *402* (6759), 276–279. https://doi.org/10.1038/46248.

(135) Férey, G. Hybrid Porous Solids: Past, Present, Future. *Chem Soc Rev* **2008**, *37* (1), 191–214. https://doi.org/10.1039/B618320B.

(136) Tranchemontagne, D. J.; Mendoza-Cortés, J. L.; O'Keeffe, M.; Yaghi, O. M. Secondary Building Units, Nets and Bonding in the Chemistry of Metal–

Organic Frameworks. *Chem. Soc. Rev.* **2009**, *38* (5), 1257. https://doi.org/10.1039/b817735j.

(137) Beck, J. S.; Vartuli, J. C.; Roth, W. J.; Leonowicz, M. E.; Kresge, C. T.; Schmitt, K. D.; Chu, C. T. W.; Olson, D. H.; Sheppard, E. W.; McCullen, S. B.; Higgins, J. B.; Schlenker, J. L. A New Family of Mesoporous Molecular Sieves Prepared with Liquid Crystal Templates. *J. Am. Chem. Soc.* **1992**, *114* (27), 10834–10843. https://doi.org/10.1021/ja00053a020.

(138) Rajan, A. C.; Mishra, A.; Satsangi, S.; Vaish, R.; Mizuseki, H.; Lee, K.-R.; Singh, A. K. Machine-Learning-Assisted Accurate Band Gap Predictions of Functionalized MXene. *Chem. Mater.* **2018**, *30* (12), 4031–4038. https://doi.org/10.1021/acs.chemmater.8b00686.

(139) Kunkel, C.; Schober, C.; Margraf, J. T.; Reuter, K.; Oberhofer, H. Finding the Right Bricks for Molecular Legos: A Data Mining Approach to Organic Semiconductor Design. *Chem. Mater.* **2019**, *31* (3), 969–978. https://doi.org/10.1021/acs.chemmater.8b04436.

(140) Davies, D. W.; Butler, K. T.; Walsh, A. Data-Driven Discovery of Photoactive Quaternary Oxides Using First-Principles Machine Learning. *Chem. Mater.* **2019**, *31* (18), 7221–7230. https://doi.org/10.1021/acs.chemmater.9b01519.

(141) Moghadam, P. Z.; Rogge, S. M. J.; Li, A.; Chow, C.-M.; Wieme, J.; Moharrami, N.; Aragones-Anglada, M.; Conduit, G.; Gomez-Gualdron, D. A.; Van Speybroeck, V.; Fairen-Jimenez, D. Structure-Mechanical Stability Relations of

Metal-Organic Frameworks via Machine Learning. *Matter* **2019**, *1* (1), 219–234. https://doi.org/10.1016/j.matt.2019.03.002.

(142) Chong, S.; Lee, S.; Kim, B.; Kim, J. Applications of Machine Learning in Metal-Organic Frameworks. *Coord. Chem. Rev.* **2020**, *423*, 213487. https://doi.org/10.1016/j.ccr.2020.213487.

(143) Pandharkar, R.; Hermes, M. R.; Cramer, C. J.; Gagliardi, L. Spin-State Ordering in Metal-Based Compounds Using the Localized Active Space Self-Consistent Field Method. *J. Phys. Chem. Lett.* **2019**, *10* (18), 5507–5513. https://doi.org/10.1021/acs.jpclett.9b02077.

(144) Smith, J. S.; Isayev, O.; Roitberg, A. E. ANI-1: An Extensible Neural Network Potential with DFT Accuracy at Force Field Computational Cost. *Chem. Sci.* **2017**, *8* (4), 3192–3203. https://doi.org/10.1039/C6SC05720A.

(145) Nguyen-Tuong, D.; Seeger, M.; Peters, J. Model Learning with Local Gaussian Process Regression. *Adv. Robot.* **2009**, *23* (15), 2015–2034. https://doi.org/10.1163/016918609X12529286896877.

(146) *Advanced Lectures on Machine Learning: ML Summer Schools 2003, Canberra, Australia, February 2-14, 2003 [and] Tübingen, Germany, August 4-16, 2003: Revised Lectures*; Bousquet, O., Luxburg, U. von, Rätsch, G., Eds.; Lecture notes in computer science, Lecture notes in artificial intelligence; Springer: Berlin ; New York, 2004.

(147) Jablonka, K. M.; Ongari, D.; Moosavi, S. M.; Smit, B. Big-Data Science in Porous Materials: Materials Genomics and Machine Learning. *Chem. Rev.* **2020**, *120* (16), 8066–8129. https://doi.org/10.1021/acs.chemrev.0c00004.

(148) Alvaro, M.; Carbonell, E.; Ferrer, B.; Llabrés i Xamena, F. X.; Garcia, H. Semiconductor Behavior of a Metal-Organic Framework (MOF). *Chem. - Eur. J.* **2007**, *13* (18), 5106–5112. https://doi.org/10.1002/chem.200601003.

(149) Narayan, T. C.; Miyakai, T.; Seki, S.; Dincă, M. High Charge Mobility in a Tetrathiafulvalene-Based Microporous Metal–Organic Framework. *J. Am. Chem. Soc.* **2012**, *134* (31), 12932–12935. https://doi.org/10.1021/ja3059827.

(150) Meng, Z.; Mirica, K. A. Two-Dimensional d-π Conjugated Metal-Organic Framework Based on Hexahydroxytrinaphthylene. *Nano Res.* **2021**, *14* (2), 369–375. https://doi.org/10.1007/s12274-020-2874-x.

(151) He, Y.; Cubuk, E. D.; Allendorf, M. D.; Reed, E. J. Metallic Metal–Organic Frameworks Predicted by the Combination of Machine Learning Methods and *Ab Initio* Calculations. *J. Phys. Chem. Lett.* **2018**, *9* (16), 4562–4569. https://doi.org/10.1021/acs.jpclett.8b01707.

(152) Eckhoff, M.; Behler, J. From Molecular Fragments to the Bulk: Development of a Neural Network Potential for MOF-5. *J. Chem. Theory Comput.* **2019**, *15* (6), 3793–3809. https://doi.org/10.1021/acs.jctc.8b01288.

(153) Wu, H.; Zhou, W.; Yildirim, T. Methane Sorption in Nanoporous Metal−Organic Frameworks and First-Order Phase Transition of Confined

Methane. *J. Phys. Chem. C* **2009**, *113* (7), 3029–3035. https://doi.org/10.1021/jp8103276.

(154) Mittal, R.; Chaplot, S. L.; Schober, H. Measurement of Anharmonicity of Phonons in the Negative Thermal Expansion Compound Zn(CN)2 by High Pressure Inelastic Neutron Scattering. *Appl. Phys. Lett.* **2009**, *95* (20), 201901. https://doi.org/10.1063/1.3264963.

(155) Dubbeldam, D.; Walton, K. S.; Ellis, D. E.; Snurr, R. Q. Exceptional Negative Thermal Expansion in Isoreticular Metal–Organic Frameworks. *Angew. Chem. Int. Ed.* **2007**, *46* (24), 4496–4499. https://doi.org/10.1002/anie.200700218.

(156) Peterson, V. K.; Kearley, G. J.; Wu, Y.; Ramirez-Cuesta, A. J.; Kemner, E.; Kepert, C. J. Local Vibrational Mechanism for Negative Thermal Expansion: A Combined Neutron Scattering and First-Principles Study. *Angew. Chem. Int. Ed.* **2010**, *49* (3), 585–588. https://doi.org/10.1002/anie.200903366.

(157) Kaur, H.; Sundriyal, S.; Kumar, V.; Sharma, A. L.; Kim, K.-H.; Wang, B.; Deep, A. Theoretical Prediction of Thermal and Electronic Properties of Metal-Organic Frameworks. *J. Ind. Eng. Chem.* **2019**, *80*, 136–151. https://doi.org/10.1016/j.jiec.2019.07.041.

(158) Chen, H.; Cheng, S.; Chen, D.; Jiang, Y.; Ang, E. H.; Liu, W.; Feng, Y.; Rui, X.; Yu, Y. Vanadate-Based Electrodes for Rechargeable Batteries. *Mater. Chem. Front.* **2021**, *5* (4), 1585–1609. https://doi.org/10.1039/D0QM00656D.

(159) Grajciar, L.; Heard, C. J.; Bondarenko, A. A.; Polynski, M. V.; Meeprasert, J.; Pidko, E. A.; Nachtigall, P. Towards *Operando* Computational

Modeling in Heterogeneous Catalysis. *Chem. Soc. Rev.* **2018**, *47* (22), 8307–8348. https://doi.org/10.1039/C8CS00398J.

(160) Gülsoy, Z.; Sezginel, K. B.; Uzun, A.; Keskin, S.; Yıldırım, R. Analysis of CH_4 Uptake over Metal–Organic Frameworks Using Data-Mining Tools. *ACS Comb. Sci.* **2019**, *21* (4), 257–268. https://doi.org/10.1021/acscombsci.8b00150.

(161) Lee, J.; Seko, A.; Shitara, K.; Nakayama, K.; Tanaka, I. Prediction Model of Band Gap for Inorganic Compounds by Combination of Density Functional Theory Calculations and Machine Learning Techniques. *Phys. Rev. B* **2016**, *93* (11), 115104. https://doi.org/10.1103/PhysRevB.93.115104.

(162) Pilania, G.; Gubernatis, J. E.; Lookman, T. Multi-Fidelity Machine Learning Models for Accurate Bandgap Predictions of Solids. *Comput. Mater. Sci.* **2017**, *129*, 156–163. https://doi.org/10.1016/j.commatsci.2016.12.004.

(163) Nørskov, J. K.; Bligaard, T.; Rossmeisl, J.; Christensen, C. H. Towards the Computational Design of Solid Catalysts. *Nat. Chem.* **2009**, *1* (1), 37–46. https://doi.org/10.1038/nchem.121.

(164) Noh, J.; Back, S.; Kim, J.; Jung, Y. Active Learning with Non- *Ab Initio* Input Features toward Efficient CO_2 Reduction Catalysts. *Chem. Sci.* **2018**, *9* (23), 5152–5159. https://doi.org/10.1039/C7SC03422A.

(165) Mao, X.; Tang, C.; He, T.; Wijethunge, D.; Yan, C.; Zhu, Z.; Du, A. Computational Screening of MN_4 (M = Ti–Cu) Based Metal Organic Frameworks for CO_2 Reduction Using the d-Band Centre as a Descriptor. *Nanoscale* **2020**, *12* (10), 6188–6194. https://doi.org/10.1039/C9NR09529B.

(166) Sturluson, A.; Huynh, M. T.; Kaija, A. R.; Laird, C.; Yoon, S.; Hou, F.; Feng, Z.; Wilmer, C. E.; Colón, Y. J.; Chung, Y. G.; Siderius, D. W.; Simon, C. M. The Role of Molecular Modelling and Simulation in the Discovery and Deployment of Metal-Organic Frameworks for Gas Storage and Separation. *Mol. Simul.* **2019**, *45* (14–15), 1082–1121. https://doi.org/10.1080/08927022.2019.1648809.

(167) Fernandez, M.; Trefiak, N. R.; Woo, T. K. Atomic Property Weighted Radial Distribution Functions Descriptors of Metal–Organic Frameworks for the Prediction of Gas Uptake Capacity. *J. Phys. Chem. C* **2013**, *117* (27), 14095–14105. https://doi.org/10.1021/jp404287t.

(168) Fernandez, M.; Boyd, P. G.; Daff, T. D.; Aghaji, M. Z.; Woo, T. K. Rapid and Accurate Machine Learning Recognition of High Performing Metal Organic Frameworks for CO_2 Capture. *J. Phys. Chem. Lett.* **2014**, *5* (17), 3056–3060. https://doi.org/10.1021/jz501331m.

(169) Pardakhti, M.; Moharreri, E.; Wanik, D.; Suib, S. L.; Srivastava, R. Machine Learning Using Combined Structural and Chemical Descriptors for Prediction of Methane Adsorption Performance of Metal Organic Frameworks (MOFs). *ACS Comb. Sci.* **2017**, *19* (10), 640–645. https://doi.org/10.1021/acscombsci.7b00056.

(170) Gharagheizi, F.; Tang, D.; Sholl, D. S. Selecting Adsorbents to Separate Diverse Near-Azeotropic Chemicals. *J. Phys. Chem. C* **2020**, *124* (6), 3664–3670. https://doi.org/10.1021/acs.jpcc.9b10955.

(171) Fanourgakis, G. S.; Gkagkas, K.; Tylianakis, E.; Klontzas, E.; Froudakis, G. A Robust Machine Learning Algorithm for the Prediction of Methane Adsorption in Nanoporous Materials. *J. Phys. Chem. A* **2019**, *123* (28), 6080–6087. https://doi.org/10.1021/acs.jpca.9b03290.

(172) Ma, S.; Liu, Z.-P. Machine Learning for Atomic Simulation and Activity Prediction in Heterogeneous Catalysis: Current Status and Future. *ACS Catal.* **2020**, *10* (22), 13213–13226. https://doi.org/10.1021/acscatal.0c03472.

(173) Liu, Z.-H.; Shi, T.-T.; Chen, Z.-X. Machine Learning Prediction of Monatomic Adsorption Energies with Non-First-Principles Calculated Quantities. *Chem. Phys. Lett.* **2020**, *755*, 137772. https://doi.org/10.1016/j.cplett.2020.137772.

(174) Bligaard, T.; Nørskov, J. K.; Dahl, S.; Matthiesen, J.; Christensen, C. H.; Sehested, J. The Brønsted–Evans–Polanyi Relation and the Volcano Curve in Heterogeneous Catalysis. *J. Catal.* **2004**, *224* (1), 206–217. https://doi.org/10.1016/j.jcat.2004.02.034.

(175) Peterson, A. A.; Nørskov, J. K. Activity Descriptors for CO_2 Electroreduction to Methane on Transition-Metal Catalysts. *J. Phys. Chem. Lett.* **2012**, *3* (2), 251–258. https://doi.org/10.1021/jz201461p.

(176) Cheng, T.; Huang, Y.; Xiao, H.; Goddard, W. A. Predicted Structures of the Active Sites Responsible for the Improved Reduction of Carbon Dioxide by Gold Nanoparticles. *J. Phys. Chem. Lett.* **2017**, *8* (14), 3317–3320. https://doi.org/10.1021/acs.jpclett.7b01335.

(177) Toyao, T.; Suzuki, K.; Kikuchi, S.; Takakusagi, S.; Shimizu, K.; Takigawa, I. Toward Effective Utilization of Methane: Machine Learning Prediction of Adsorption Energies on Metal Alloys. *J. Phys. Chem. C* **2018**, *122* (15), 8315–8326. https://doi.org/10.1021/acs.jpcc.7b12670.

(178) Fung, V.; Hu, G.; Ganesh, P.; Sumpter, B. G. Machine Learned Features from Density of States for Accurate Adsorption Energy Prediction. *Nat. Commun.* **2021**, *12* (1), 88. https://doi.org/10.1038/s41467-020-20342-6.

(179) Chen, B. W. J.; Xu, L.; Mavrikakis, M. Computational Methods in Heterogeneous Catalysis. *Chem. Rev.* **2021**, *121* (2), 1007–1048. https://doi.org/10.1021/acs.chemrev.0c01060.

(180) Guan, Y.; Chaffart, D.; Liu, G.; Tan, Z.; Zhang, D.; Wang, Y.; Li, J.; Ricardez-Sandoval, L. Machine Learning in Solid Heterogeneous Catalysis: Recent Developments, Challenges and Perspectives. *Chem. Eng. Sci.* **2022**, *248*, 117224. https://doi.org/10.1016/j.ces.2021.117224.

(181) Schlexer Lamoureux, P.; Winther, K. T.; Garrido Torres, J. A.; Streibel, V.; Zhao, M.; Bajdich, M.; Abild-Pedersen, F.; Bligaard, T. Machine Learning for Computational Heterogeneous Catalysis. *ChemCatChem* **2019**, *11* (16), 3581–3601. https://doi.org/10.1002/cctc.201900595.

(182) Xu, J.; Cao, X.-M.; Hu, P. Perspective on Computational Reaction Prediction Using Machine Learning Methods in Heterogeneous Catalysis. *Phys. Chem. Chem. Phys.* **2021**, *23* (19), 11155–11179. https://doi.org/10.1039/D1CP01349A.

(183) Mamun, O.; Winther, K. T.; Boes, J. R.; Bligaard, T. High-Throughput Calculations of Catalytic Properties of Bimetallic Alloy Surfaces. *Sci. Data* **2019**, *6* (1), 76. https://doi.org/10.1038/s41597-019-0080-z.

(184) *NIST Chemistry WebBook*; Linstrom, P. J., Mallard, W. G., Eds.; NIST Standard Reference Database; Vol. 69.

(185) Jain, A.; Ong, S. P.; Hautier, G.; Chen, W.; Richards, W. D.; Dacek, S.; Cholia, S.; Gunter, D.; Skinner, D.; Ceder, G.; Persson, K. A. Commentary: The Materials Project: A Materials Genome Approach to Accelerating Materials Innovation. *APL Mater.* **2013**, *1* (1), 011002. https://doi.org/10.1063/1.4812323.

(186) Guo, C.; Berkhahn, F. Entity Embeddings of Categorical Variables. arXiv April 22, 2016. http://arxiv.org/abs/1604.06737 (accessed 2022-09-28).

(187) Dorogush, A. V.; Ershov, V.; Gulin, A. CatBoost: Gradient Boosting with Categorical Features Support. arXiv October 24, 2018. http://arxiv.org/abs/1810.11363 (accessed 2022-09-28).

(188) Giordano, L.; Akkiraju, K.; Jacobs, R.; Vivona, D.; Morgan, D.; Shao-Horn, Y. Electronic Structure-Based Descriptors for Oxide Properties and Functions. *Acc. Chem. Res.* **2022**, *55* (3), 298–308. https://doi.org/10.1021/acs.accounts.1c00509.

(189) Goldsmith, B. R.; Esterhuizen, J.; Liu, J.; Bartel, C. J.; Sutton, C. Machine Learning for Heterogeneous Catalyst Design and Discovery. *AIChE J.* **2018**, *64* (7), 2311–2323. https://doi.org/10.1002/aic.16198.

(190) Gusarov, S.; Stoyanov, S. R.; Siahrostami, S. Development of Fukui Function Based Descriptors for a Machine Learning Study of CO $_2$ Reduction. *J. Phys. Chem. C* **2020**, *124* (18), 10079–10084. https://doi.org/10.1021/acs.jpcc.0c03101.

(191) Esterhuizen, J. A.; Goldsmith, B. R.; Linic, S. Theory-Guided Machine Learning Finds Geometric Structure-Property Relationships for Chemisorption on Subsurface Alloys. *Chem* **2020**, *6* (11), 3100–3117. https://doi.org/10.1016/j.chempr.2020.09.001.

(192) Hancock, J. T.; Khoshgoftaar, T. M. Survey on Categorical Data for Neural Networks. *J. Big Data* **2020**, *7* (1), 28. https://doi.org/10.1186/s40537-020-00305-w.

(193) Grinsztajn, L.; Oyallon, E.; Varoquaux, G. Why Do Tree-Based Models Still Outperform Deep Learning on Tabular Data? arXiv July 18, 2022. http://arxiv.org/abs/2207.08815 (accessed 2022-10-12).

(194) Chen, T.; Guestrin, C. XGBoost: A Scalable Tree Boosting System. In *Proceedings of the 22nd ACM SIGKDD International Conference on Knowledge Discovery and Data Mining*; 2016; pp 785–794. https://doi.org/10.1145/2939672.2939785.

(195) Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; Vanderplas, J.; Passos, A.; Cournapeau, D. Scikit-Learn: Machine Learning in Python. *Mach. Learn. PYTHON* 6.

(196) Ke, G.; Meng, Q.; Finley, T.; Wang, T.; Chen, W.; Ma, W.; Ye, Q.; Liu, T.-Y. LightGBM: A Highly Efficient Gradient Boosting Decision Tree. 9.

(197) Akiba, T.; Sano, S.; Yanase, T.; Ohta, T.; Koyama, M. Optuna: A Next-Generation Hyperparameter Optimization Framework. In *Proceedings of the 25th ACM SIGKDD International Conference on Knowledge Discovery & Data Mining*; ACM: Anchorage AK USA, 2019; pp 2623–2631. https://doi.org/10.1145/3292500.3330701.

(198) Kapoor, S.; Narayanan, A. Leakage and the Reproducibility Crisis in ML-Based Science. arXiv July 14, 2022. http://arxiv.org/abs/2207.07048 (accessed 2022-09-28).

(199) Artrith, N.; Butler, K. T.; Coudert, F.-X.; Han, S.; Isayev, O.; Jain, A.; Walsh, A. Best Practices in Machine Learning for Chemistry. *Nat. Chem.* **2021**, *13* (6), 505–508. https://doi.org/10.1038/s41557-021-00716-z.

(200) Vishwakarma, G.; Sonpal, A.; Hachmann, J. Metrics for Benchmarking and Uncertainty Quantification: Quality, Applicability, and Best Practices for Machine Learning in Chemistry. *Trends Chem.* **2021**, *3* (2), 146–156. https://doi.org/10.1016/j.trechm.2020.12.004.

(201) Lundberg, S.; Lee, S.-I. A Unified Approach to Interpreting Model Predictions. arXiv November 24, 2017. http://arxiv.org/abs/1705.07874 (accessed 2022-09-28).

(202) Praveen, C. S.; Comas-Vives, A. Design of an Accurate Machine Learning Algorithm to Predict the Binding Energies of Several Adsorbates on Multiple Sites

of Metal Surfaces. *ChemCatChem* **2020**, *12* (18), 4611–4617. https://doi.org/10.1002/cctc.202000517.

(203) Akinola, J.; Campbell, C. T.; Singh, N. Effects of Solvents on Adsorption Energies: A General Bond-Additivity Model. *J. Phys. Chem. C* **2021**, *125* (44), 24371–24380. https://doi.org/10.1021/acs.jpcc.1c06781.