TOWARD REALISTIC MODELING OF CATALYTIC SURFACES: FROM FIRST PRINCIPLES TO MACHINE LEARNING

Robert B. Wexler

A DISSERTATION

in

Chemistry

Presented to the Faculties of the University of Pennsylvania

in

Partial Fulfillment of the Requirements for the
Degree of Doctor of Philosophy

2019

Supervisor of Dissertation

Dr. Andrew M. Rappe
Blanchard Professor of Chemistry, and
Professor of Materials Science and Engineering

Graduate Group Chairperson

Dr. David W. Christianson
Roy and Diana Vagelos Professor in Chemistry and Chemical Biology

Dissertation Committee
Dr. Jeffrey G. Saven, Professor of Chemistry
Dr. Christopher B. Murray, Richard Perry University Professor of Chemistry and Materials Science and Engineering
Dr. Vaclav Vitek, Harold Pender Professor of Materials Science and Engineering, and Mechanical Engineering and Applied Mechanics

TOWARD REALISTIC MODELING OF CATALYTIC SURFACES: FROM FIRST PRINCIPLES TO MACHINE LEARNING

To my family: Shani, Mom, Dad, Jourdan, Juliah, and Andrea.

ACKNOWLEDGEMENTS

I would like to thank Prof. Andrew M. Rappe for motivating me to set lofty goals, being understanding in difficult times, encouraging me to think creatively, and providing me with opportunities to develop a well-rounded set of research skills.

I would also like to thank Dr. J. Mark P. Martirez for teaching me about *ab initio* thermodynamics and electrocatalysis, answering my (many) questions, mentoring me in my first graduate research project, and advocating for me.

I would like to express my appreciation to Prof. G. Charles Dismukes for sharing his unique perspective on electrocatalyst design, supporting me in my job applications, and imparting helpful career and life advice. Also, to Dr. Anders B. Laursen for stimulating conversations and our fruitful collaboration.

I would like to offer my thanks to the members of the Rappe group, past and present, for countless insightful discussions, their willingness to assist me, and the lasting friendships. To Mr. Tian Qiu for helping me develop *ab initio* grand canonical Monte Carlo and giving of your time so generously. To Dr. Arvin Kakekhani for deepening my understanding of catalysis.

Finally, I am grateful for the financial and computational support given by the following agencies: the US Department of Energy, the Office of Naval Research, the US Department of Defense High Performance Computing Modernization Program, and the National Energy Research Scientific Computing Center.

ABSTRACT

TOWARD REALISTIC MODELING OF CATALYTIC SURFACES: FROM FIRST PRINCIPLES TO MACHINE LEARNING

Robert B. Wexler

Dr. Andrew M. Rappe

Computational catalyst design has the potential to revolutionize the energy and chemical industries by alleviating their reliance on fossil fuels, precious metals, and toxic elements. Despite recent advances in understanding catalytic trends, *e.g.* the chemisorption scaling relations and the *d*-band model, the description of catalytic surfaces has, for the most part, been far from realistic. It is well known that surfaces can undergo reconstruction where the structure and composition of the surface differs from that of the bulk and the nature of this reconstruction depends on the temperature, pressure, and chemical potentials of the elements in the system. Since catalytic transformations, *i.e.* bond breaking and formation, occur at the surface, an accurate picture of surface structure and composition is vital. In this thesis, we apply and develop state-of-the-art computational methods for studying the reconstruction of catalytic surfaces and investigate the effect of surface reconstruction on catalysis. First, we show using *ab initio* thermodynamics that the surfaces of nickel phosphide catalysts for the hydrogen evolution reaction (HER) are P-enriched, which was not previously considered in computational studies. Building on this discovery, we reevaluate the HER mechanism on Ni_2P and Ni_5P_4 and find that P sites are the key to their catalytic activities. While the P sites on Ni_2P are highly active toward the HER, they are not stable at conditions suitable for commercial electrolyzers. Under

these conditions, the stable surface of Ni_2P binds hydrogen too strongly at Ni sites. We demonstrate that these Ni sites can be activated by doping the surface of Ni_2P with S, Se, and Te. Additionally, using tree-based machine learning methods, we reveal that nonmetal dopants induce a chemical pressure-like effect on the Ni sites, changing their reactivity through compression and expansion. Finally, we develop a software package for *ab initio* grand canonical Monte Carlo that automatically predicts surface phase diagrams. The results presented herein provide strong motivation and a methodological foundation for moving toward more realistic modeling of heterogeneous catalysts.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF ILLUSTRATIONS

PREFACE

The materials in chapter 3 appeared in R. B. Wexler, J. M. P. Martirez, and A. M. Rappe, "Stable Phosphorus-Enriched (0001) Surfaces of Nickel Phosphides", *Chem. Mater.*, **2016**, *28* (15), pp 5365–5372. Copyright © 2016 American Chemical Society.

The materials in chapter 4 appeared in R. B. Wexler, J. M. P. Martirez, and A. M. Rappe, "Active Role of Phosphorus in the Hydrogen Evolving Activity of Nickel Phosphide (0001) Surfaces", *ACS Catal.*, **2017**, *7* (11), pp 7718–7725. Copyright © 2017 American Chemical Society.

The materials in chapter 5 appeared in R. B. Wexler, J. M. P. Martirez, and A. M. Rappe, "Chemical Pressure-Driven Enhancement of the Hydrogen Evolving Activity of Ni_2P from Nonmetal Surface Doping Interpreted via Machine Learning", *J. Am. Chem. Soc.*, **2018**, *140* (13), pp 4678–4683. Copyright © 2018 American Chemical Society.

The materials in chapter 6 appeared in R. B. Wexler, T. Qiu, and A. M. Rappe, "Automatic Prediction of Surface Phase Diagrams Using *Ab Initio* Grand Canonical Monte Carlo", *J. Phys. Chem. C*, **2019**, *123* (4), pp 2321–2328. Copyright © 2019 American Chemical Society.

CHAPTER 1 : Introduction

Heterogeneous catalysis is of tremendous importance to a number of industries, including the chemical and energy industries, and, more broadly, our global economy. (6–9) It has implications in the synthesis of fertilizers that sustain our food supply, (10–12) materials that make up our clothing, (13–15) and the processing of chemicals that fuel our vehicles. (16–18) In the last few decades, it has also been applied to the pursuit of sustainable, eco-friendly energy solutions such as the remediation of atmospheric CO_2 (19–21) and the renewable production of H_2, (22–24) which finds application in many industrial processes, via the electrolysis of water. Regarding the latter, there has been a great deal of effort to replace Pt, the prototypical water splitting electrocatalyst, with elements or compounds that are more abundant and inexpensive. (23; 25) Of the many proposed alternatives, one of the most popular in recent years has been the transition metal phosphides, specifically those based on Fe, (26–28) Co, (29–31) and Ni. (32–34) Nickel phosphides specifically have been the subject of numerous investigations, both experimental (32; 34; 35) and theoretical, (27; 33; 36–40) that aim to provide a deep mechanistic understanding of their activity toward the hydrogen evolution reaction (HER) and subsequently to identify materials design principles for engineering their catalytic properties. Many of the computational studies of the HER on nickel phosphides (27; 33; 36) and, more generally, of heterogeneous catalysis on complex materials, however, make unrealistic assumptions about the catalyst surface, namely that is does not undergo any sort of stabilizing reconstruction.

Surface reconstruction is a process in which the structure and composition of the surface, and sometimes layers near the surface, differ from that of the bulk. Early

studies on Si show that, despite its crystallographic simplicity in the bulk, it can acquire a variety of surface reconstructions with different periodicities (41–43) including one with a remarkable dimer-adatom-stacking-fault geometry (see Fig. 1.1a). (44–46) Another classic example of surface reconstruction is that of Au(110), which has missing rows of Au atoms along [100] (see Fig. 1.1b). (47; 48) For materials with two or more elements and complicated crystal structures, surface reconstruction behavior becomes even more exotic, *e.g.* the formation of TiO_x double layers on $ATiO_3$ perovskites (A being Ba, Sr, or Pb). (49–51) The primary driving forces of surface reconstruction on semiconductors, metals, and ionic solids differ substantially where semiconductors aim to saturate surface atom coordination, metals seek to increase the surface packing efficiency, and ionic crystals prefer to minimize the electrostatic energy by passivating surface dipoles. Notwithstanding this evidence, the computational catalysis community still has yet to fully embrace the prevalence of surface reconstruction and its potentially dramatic impact on chemical reactivity.

One of the reasons for this is that, until recently, the computational methods for predicting surface phase diagrams lacked the requisite combination of first principles accuracy, computational efficiency, and completeness in terms of sampling the phase space of surface reconstructions. Today, the standard technique for computing surface phase diagrams is *ab initio* thermodynamics, which unites first principles density functional theory (DFT) calculations, the statistical mechanics of vibrations, experimental thermochemistry data, and the relationship between thermodynamic quantities such as the Gibbs free energy, temperature, pressure, and chemical potential. (52; 53) This method can be further extended to deal with aqueous electrochemical environments and used to construct surface phase Pourbaix diagrams, which are plotted as a function of pH, activity, and the applied electrode potential. (54) While *ab initio* thermo-

dynamics has found great success in reproducing the experimentally observed surface structures for a diverse set of materials such as Ag(111), (55; 56) RuO_2(110), (57) $LiNbO_3$(001), (58) and $LaMnO_3$(001) (54) to name a few, it is limited by the fact that the catalog of surfaces used to construct the surface phase diagram is prepared manually based on chemical intuition. It therefore lacks a systematic way to control the completeness with which phase space is sampled and may lead to incomplete or, worse yet, incorrect prediction of surface phase diagrams due to human bias. Recently, there have been a number of developments in surface crystal structure prediction involving metaheuristic optimization algorithms, *e.g.* genetic/evolutionary (59–62) and particle swarm techniques, (63–65) as well as machine learning models, like Gaussian processes, (66) that can be used to more efficiently traverse phase space. These more advanced methods have proven quite successful in identifying new surface phases with exciting properties, however, they are somewhat black box and do not offer chemical and physical insight into the natural stochastic evolution of a surface, which ideally one might desire.

In this thesis, we apply *ab initio* thermodynamics to the study of nickel phosphide electrocatalysts for the HER (38; 39; 67) and develop a new methodology for the automatic prediction of surface phase diagrams based on *ab initio* grand canonical Monte Carlo. (68) In Chapter 2, we lay the theoretical groundwork for DFT, *ab initio* thermodynamics, grand canonical Monte Carlo, and tree-based machine learning methods. In Chapter 3, we explore the reconstructions of the (0001) surface of Ni_2P and Ni_5P_4 in equilibrium with bulk Ni and P. (67) In Chapter 4, we recalculate these phase diagrams using *ab initio* aqueous thermodynamics and use them to model the HER under acidic conditions. (38) In Chapter 5, we devise nonmetal surface doping schemes for improving the activity of Ni_2P(0001) and identify catalytic descriptors

using random forest regression. (39) Finally, in Chapter 6, we develop a new software for surface crystal structure prediction using *ab initio* grand canonical Monte Carlo and reproduce the experimental surface reconstructions of Ag(111) as a proof of concept. (68) More detailed summaries of Chapters 3 through 6 are presented below.

In Chapter 3, we explore the reconstructions of both $Ni_2P(0001)$ and $Ni_5P_4(000\pm1)$ surfaces with first principles DFT. Most of the stable terminations under realistic synthesis conditions are determined to be P-rich on both materials. A P-covered reconstruction of the Ni_3P_2 termination of $Ni_2P(0001)$ is found to be most stable, consistent with the current literature. By contrast, the most energetically favorable surfaces of Ni_5P_4 are found to be the Ni_3P_3 and Ni_4P_3 bulk-derived terminations with P-adatoms. The preferred excess P binding sites and their energies are identified on each surface. We find that the P_3 site which is present on Ni_5P_4, and the Ni_3 site, which is present on both Ni_2P and Ni_5P_4, strongly bind excess P. Additionally, we predict the presence of stable P_n ($n = 2, 4$) agglomerates on Ni_5P_4 at the P_3-hollow and Ni-Ni bridge sites. This chapter highlights the importance of considering the aggregation behavior of non-metal components in predicting the surface reconstruction of transition metal compounds. (67)

In Chapter 4, we explore, through DFT with thermodynamics, the aqueous reconstructions of $Ni_2P(0001)$ and $Ni_5P_4(0001)/(000\pm1)$ and find that the surface P content on $Ni_2P(0001)$ depends on the applied potential, which has not been considered previously. At -0.21 V $\geq U \geq$ -0.36 V *vs.* the standard hydrogen electrode and pH = 0, a PH_x-enriched Ni_3P_2 termination of $Ni_2P(0001)$ is found to be most stable, consistent with its P-rich ultra-high-vacuum reconstructions. Above and below this potential range, the stoichiometric Ni_3P_2 surface is instead passivated by H at the Ni_3-hollow sites. On the other hand, $Ni_5P_4(000\bar{1})$ does not favor additional

P. Instead, the Ni_4P_3 bulk termination of $Ni_5P_4(000\bar{1})$ is passivated by H at both the Ni_3 and P_3-hollow sites. We also found that the most HER-active surfaces are Ni_3P_2+P+(7/3)H of $Ni_2P(0001)$ and Ni_4P_3+4H of $Ni_5P_4(000\bar{1})$ due to weak H adsorption at P catalytic sites, in contrast with other computational investigations that propose Ni as or part of the active site. By looking at viable catalytic cycles for HER on the stable reconstructed surfaces, and calculating the reaction free energies of the associated elementary steps, we calculate that the overpotential on the Ni_4P_3+4H surface of $Ni_5P_4(000\bar{1})$ (-0.16 V) is lower than that of the Ni_3P_2+P+(7/3)H surface of $Ni_2P(0001)$ (-0.21 V). This is due to the abundance of P_3-hollow sites on Ni_5P_4 and the limited surface stability of the P-enriched $Ni_2P(0001)$ surface phase. The trend in the calculated overpotentials explains why the nickel phosphides studied here perform almost as well as Pt, and why Ni_5P_4 is superior to Ni_2P toward HER, as is found in the experimental literature. This chapter emphasizes the importance of considering aqueous surface stability in predicting the HER-active sites, mechanism, and overpotential, and highlights the primary role of P in HER catalysis on transition metal phosphides. (38)

In Chapter 5, we investigate the effect of surface nonmetal doping on the HER activity of the Ni_3P_2 termination of $Ni_2P(0001)$, which is stable at modest electrochemical conditions. Using DFT calculations, we find that both $2p$ nonmetals and heavier chalcogens provide nearly thermoneutral H adsorption at moderate surface doping concentrations. We also find, however, that only chalcogen substitution for surface P is exergonic. For intermediate surface concentrations of S, the free energy of H adsorption at the Ni_3-hollow site is -0.11 eV, which is significantly more thermoneutral than the undoped surface (-0.45 eV). We use the regularized random forest machine learning algorithm to discover the relative importance of structure and

charge descriptors, extracted from the DFT calculations, in determining the HER activity of $Ni_2P(0001)$ under different doping concentrations. We discover that the Ni-Ni bond length is the most important descriptor of HER activity, which suggests that the nonmetal dopants induce a chemical pressure-like effect on the Ni_3-hollow site, changing its reactivity through compression and expansion. (39)

To overcome the limitations of *ab initio* thermodynamics and automate the discovery of realistic surfaces, we combine density functional theory and grand canonical Monte Carlo (GCMC) into "*ab initio* GCMC." Chapter 6 presents the successful application of *ab initio* GCMC to the study of oxide overlayers on Ag(111), which, for many years, mystified experts in surface science and catalysis. Specifically, we report that *ab initio* GCMC is able to reproduce the surface phase diagram of Ag(111) with no preconceived notions about the system. Using nonlinear, random forest regression, we discover that Ag coordination number with O and the surface O-Ag-O bond angles are good descriptors of the surface energy. Additionally, using the composition-structure evolution histories produced by *ab initio* GCMC, we deduce a mechanism for the formation of oxide overlayers based on the Ag_3O_4 pyramid motif that is common to many reconstructions of Ag(111). In conclusion, *ab initio* GCMC is a promising tool for the discovery of realistic surfaces that can then be used to study phenomena on complex surfaces such as heterogeneous catalysis and materials growth, enabling reliable and insightful interpretations of experiments. (68)

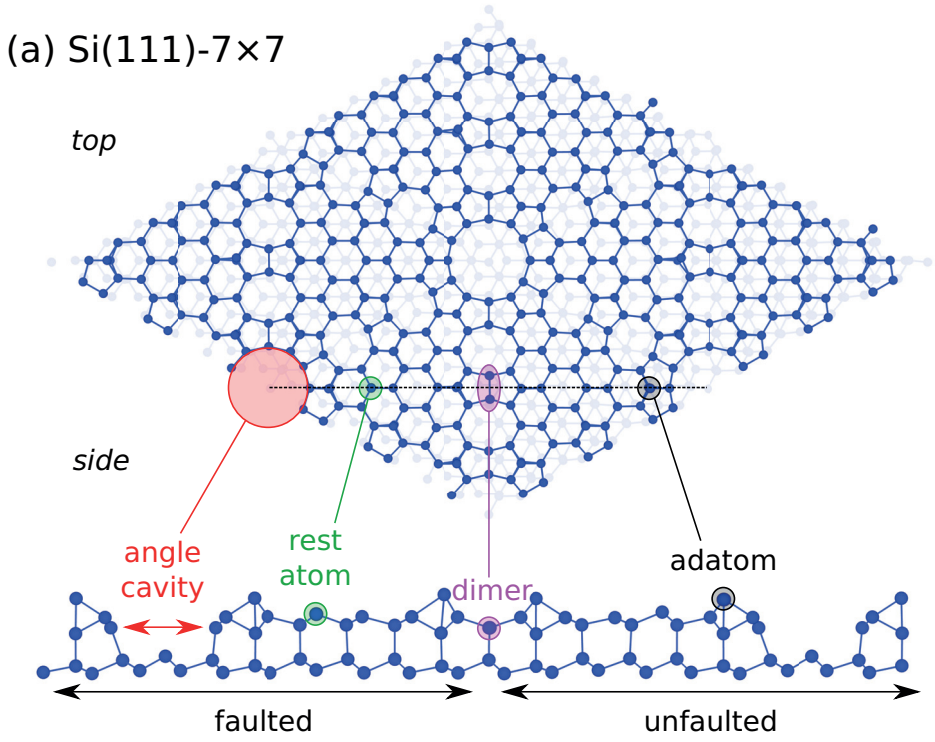

Figure 1.1: Classic examples of surface reconstruction, (a) Si(111)-7×7 and (b) Au(110)-1×2. Blue and yellow spheres correspond to silicon and gold atoms, respectively. Si(111)-7×7 has angle cavities (red), rest atoms (green), dimers (magenta), and adatoms (black). The side view is obtained by cutting the top view along the dashed line and then rotating about it by 90°. Half of the surface unit cell of Si(111)-7×7 has a stacking fault arrangement whereas the other half does not. Au(110)-1×2 has missing rows along [100] that repeat every two surface unit cells.

CHAPTER 2 : Methodology

This thesis primarily consists of thermodynamic analyses of surface chemical reactions, namely surface reconstruction and heterogeneous catalysis. The thermodynamic properties of these processes, such as surface and adsorption free energies, can be readily computed given the total energies of the reactants and products and the forces acting on the atoms. The most accurate approaches for calculating total energies and forces are based on quantum mechanics. The quantum mechanical operator for the total energy is the Hamiltonian, which takes the following form for a polyatomic system

$$\hat{H} = \sum_{I=1}^{N} -\frac{\hbar^2}{2M_I}\nabla_I^2 + \sum_{I>J} \frac{Z_I Z_J e^2}{|\mathbf{R}_I - \mathbf{R}_J|} + \sum_{i=1}^{n} -\frac{\hbar^2}{2m_i}\nabla_i^2 + \sum_{i>j} \frac{e^2}{|\mathbf{r}_i - \mathbf{r}_j|} - \sum_{i=1}^{n}\sum_{I=1}^{N} \frac{Z_I e^2}{|\mathbf{r}_i - \mathbf{R}_I|} \tag{2.1}$$

where I/J and i/j are the nucleus and electron indices, N and n are the number of nuclei and electrons, $\hbar = h/2\pi$ is the reduced Planck constant, M and m are the nucleus and electron mass, ∇^2 is the Laplace operator, Z is the nucleus charge in units of the electron charge e, and $\mathbf{R}$ and $\mathbf{r}$ are the positions of the nuclei and electrons. The first and second terms correspond to the kinetic energy and electrostatic repulsion of the nuclei, the third and fourth terms correspond to the kinetic energy ($\hat{T}_{ee}$) and electrostatic repulsion ($\hat{V}_{ee}$) of the electrons (where $\hat{F} = \hat{T}_{ee} + \hat{V}_{ee}$), and the fifth term corresponds to the electrostatic attraction between the nuclei and electrons ($\hat{V}$). As a first step to simplify Eq. 2.1, we make the Born-Oppenheimer approximation, which allows us to neglect the kinetic energy of the nuclei because they are practically stationary relative to the time scale of the electron motions. (69) By doing so, the Hamiltonian now depends only parametrically on the positions of the nuclei. What

remains is to solve the electronic Schrödinger equation (70)

$$\hat{H}\Psi\left(\mathbf{r};\mathbf{R}\right)=\left(\hat{F}+\hat{V}\right)\Psi\left(\mathbf{r};\mathbf{R}\right)=E\left(\mathbf{R}\right)\Psi\left(\mathbf{r};\mathbf{R}\right) \tag{2.2}$$

where Ψ is the electronic wave function and E is the total energy.

2.1. Density functional theory

One of the most powerful approaches to solve Eq. 2.2 is density functional theory (DFT). (71; 72) The key intuition of DFT is that the potential of the nuclei acts on only one electron at a time and therefore all one needs is the electron density to determine the potential energy. Based on this insight, we can write the potential as a sum over the potentials acting on each electron, *i.e.*

$$\hat{V}=\sum_{i}V\left(\mathbf{r}_{i}\right) \tag{2.3}$$

The expectation value of the potential is

$$\begin{aligned}\left\langle \Psi\left|\hat{V}\right|\Psi\right\rangle &= \int d\mathbf{r}_{1}\cdots d\mathbf{r}_{n}\Psi^{*}\left(\mathbf{r}_{1},\cdots,\mathbf{r}_{n}\right)\hat{V}\Psi\left(\mathbf{r}_{1},\cdots,\mathbf{r}_{n}\right)\\ &= \sum_{j}\int d\mathbf{r}_{1}\cdots d\mathbf{r}_{j}\cdots d\mathbf{r}_{n}\Psi^{*}\left(\mathbf{r}_{1},...,\mathbf{r}_{j},...,\mathbf{r}_{n}\right)V\left(\mathbf{r}_{j}\right)\Psi\left(\mathbf{r}_{1},...,\mathbf{r}_{j},...,\mathbf{r}_{n}\right)\\ &= \sum_{j}\int d\mathbf{r}_{j}\rho_{j}\left(\mathbf{r}_{j}\right)V\left(\mathbf{r}_{j}\right)=\int d\mathbf{r}\rho\left(\mathbf{r}\right)V\left(\mathbf{r}\right)\end{aligned} \tag{2.4}$$

where $\rho\left(\mathbf{r}\right)$ is the electron density.

2.1.1. Hohenberg-Kohn theorem

Using Eqs. 2.2 and 2.4, we can now prove the Hohenberg-Kohn (HK) theorem, which states that the potential is uniquely determined by the electron density. (71) Given two nondegenerate potentials $\hat{V}_1$ and $\hat{V}_2$, whose corresponding Ψ_1 and Ψ_2 are normalized, we know from Eq. 2.2 that

$$\begin{aligned} E_1 &= \int \Psi_1^* \left(\hat{F} + \hat{V}_1\right) \Psi_1 \\ E_2 &= \int \Psi_2^* \left(\hat{F} + \hat{V}_2\right) \Psi_2 \end{aligned} \tag{2.5}$$

where we have omitted differentials and position dependence for clarity. Additionally, we know from the variational principle that

$$\begin{aligned} \int \Psi_1^* \left(\hat{F} + \hat{V}_2\right) \Psi_1 &> E_2 \\ \int \Psi_2^* \left(\hat{F} + \hat{V}_1\right) \Psi_2 &> E_1 \end{aligned} \tag{2.6}$$

Inserting Eq. 2.5 into 2.6 and summing the two inequalities, we arrive at

$$\int \Psi_1^* \hat{F} \Psi_1 + \int \rho_1 \hat{V}_2 + \int \Psi_2^* \hat{F} \Psi_2 + \int \rho_2 \hat{V}_1 > \int \Psi_1^* \hat{F} \Psi_1 + \int \rho_1 \hat{V}_1 + \int \Psi_2^* \hat{F} \Psi_2 + \int \rho_2 \hat{V}_2 \tag{2.7}$$

If we assume that the potential is not uniquely determined by the electron density, in other words if $\rho_1 = \rho_2$, then Eq. 2.7 reduces to $0 > 0$, which is clearly false, thus proving the HK theorem by *reductio ad absurdum*. This theorem makes the remarkable statement that the total energy can, in principle, be written as a functional of the electron density, a real and experimentally measurable function of 3 dimensions, instead of the electronic wave function, a complex function of $3N$ dimensions.

2.1.2. Kohn-Sham equations

Having proven the HK theorem, we can write a total energy functional with the following general form

$$E\left[\rho\left(\mathbf{r}\right)\right] = \int d\mathbf{r} V\left(\mathbf{r}\right)\rho\left(\mathbf{r}\right) + F\left[\rho\left(\mathbf{r}\right)\right] \tag{2.8}$$

where F contains the electrostatic repulsion of the electrons, their kinetic energy, and exchange and correlation effects. In practice, electron-electron repulsion is included via classical electrostatics in the so-called "Hartree" energy

$$E_{\mathrm{H}}\left[\rho\left(\mathbf{r}\right)\right] = \frac{1}{2}\int\int d\mathbf{r}d\mathbf{r}'\frac{\rho\left(\mathbf{r}\right)\rho\left(\mathbf{r}'\right)}{\left|\mathbf{r}-\mathbf{r}'\right|} \tag{2.9}$$

For the kinetic energy of the electrons, we introduce a set of noninteracting, single-electron "Kohn-Sham" wave functions $\{\phi_i\left(\mathbf{r}\right)\}$ such that

$$T_{\mathrm{KS}}\left[\rho\left(\mathbf{r}\right)\right] = \sum_{i=1}^{n} -\frac{\hbar^2}{2m}\int d\mathbf{r}\phi_i^*\left(\mathbf{r}\right)\nabla^2\phi_i\left(\mathbf{r}\right) \tag{2.10}$$

where

$$\rho\left(\mathbf{r}\right) = \sum_{i}^{n}\phi_i^*\left(\mathbf{r}\right)\phi_i\left(\mathbf{r}\right) \tag{2.11}$$

While the introduction of KS orbitals runs counter to the spirit of the HK theorem, state-of-the-art kinetic energy density functions tend to fail for materials with highly localized electron densities. (71; 73–75) Inserting Eqs. 2.9 and 2.10 into 2.8 yields

$$\begin{aligned} E\left[\rho\left(\mathbf{r}\right)\right] = \int d\mathbf{r} V\left(\mathbf{r}\right)\rho\left(\mathbf{r}\right) + \frac{1}{2}\int\int d\mathbf{r}d\mathbf{r}'\frac{\rho\left(\mathbf{r}\right)\rho\left(\mathbf{r}'\right)}{\left|\mathbf{r}-\mathbf{r}'\right|} \\ + \sum_{i=1}^{n} -\frac{\hbar^2}{2m}\int d\mathbf{r}\phi_i^*\left(\mathbf{r}\right)\nabla^2\phi_i\left(\mathbf{r}\right) + E_{\mathrm{xc}}\left[\rho\left(\mathbf{r}\right)\right] \end{aligned} \tag{2.12}$$

where E_{xc} is the exchange-correlation energy, which contains electron exchange and correlation effects as well as corrections for the self-interaction error (76) and the kinetic energy of interacting electrons. It is customary to rewrite the exchange-correlation functional as

$$E_{\mathrm{xc}}\left[\rho\left(\mathbf{r}\right)\right]=\int d\mathbf{r}\epsilon_{\mathrm{xc}}\left[\rho\left(\mathbf{r}\right)\right]\rho\left(\mathbf{r}\right) \tag{2.13}$$

In order to find the ground state total energy, we apply the variational principle by taking the functional derivative of Eq. 2.12 with respect to the KS orbitals

$$\frac{\partial E\left[\rho\left(\mathbf{r}\right)\right]}{\partial\phi_{i}^{*}\left(\mathbf{r}''\right)}=0 \tag{2.14}$$

The first step in evaluating Eq. 2.14 is taking the derivative of the electron density

$$\frac{\partial\rho\left(\mathbf{r}\right)}{\partial\phi_{i}^{*}\left(\mathbf{r}''\right)}=\frac{\partial}{\partial\phi_{i}^{*}\left(\mathbf{r}''\right)}\sum_{j=1}^{n}\phi_{j}^{*}\left(\mathbf{r}\right)\phi_{j}\left(\mathbf{r}\right)=\delta\left(\mathbf{r}-\mathbf{r}''\right)\phi_{i}\left(\mathbf{r}\right) \tag{2.15}$$

Upon inserting Eqs. 2.12, 2.13, and 2.15 into 2.14 and integrating, we obtain

$$\begin{aligned}\frac{\partial E\left[\rho\left(\mathbf{r}\right)\right]}{\partial\phi_{i}^{*}\left(\mathbf{r}''\right)}=-\frac{\hbar^{2}}{2m}\nabla^{2}\phi_{i}\left(\mathbf{r}''\right)+V\left(\mathbf{r}''\right)\phi_{i}\left(\mathbf{r}''\right)+\int d\mathbf{r}'\frac{\rho\left(\mathbf{r}'\right)}{\left|\mathbf{r}''-\mathbf{r}'\right|}\phi\left(\mathbf{r}''\right)\\ +\left(\frac{\partial\epsilon_{\mathrm{xc}}\left(\mathbf{r}''\right)}{\partial\rho\left(\mathbf{r}''\right)}\rho\left(\mathbf{r}''\right)+\epsilon_{\mathrm{xc}}\left(\mathbf{r}''\right)\right)\phi\left(\mathbf{r}''\right)-\lambda_{i}\phi\left(\mathbf{r}''\right)=0\end{aligned} \tag{2.16}$$

where λ is a Lagrange multiplier that subjects the KS orbitals to an orthonormality constraint. Eq. 2.16 can be recast as

$$\left[-\frac{\hbar^{2}}{2m}\nabla^{2}+V\left(\mathbf{r}\right)+V_{\mathrm{H}}\left(\mathbf{r}\right)+V_{\mathrm{xc}}\left(\mathbf{r}\right)\right]\phi_{i}\left(\mathbf{r}\right)=\lambda_{i}\phi_{i}\left(\mathbf{r}\right) \tag{2.17}$$

where

$$V_{\mathrm{H}}(\mathbf{r}) = \int d\mathbf{r}' \frac{\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} \tag{2.18}$$

is the Hartree potential and

$$V_{\mathrm{xc}}(\mathbf{r}) = \frac{\partial \epsilon_{\mathrm{xc}}(\mathbf{r})}{\partial \rho(\mathbf{r})} \rho(\mathbf{r}) + \epsilon_{\mathrm{xc}}(\mathbf{r}) \tag{2.19}$$

is the exchange-correlation potential. Eq. 2.17 is the familiar form of the KS equations, which are traditionally solved using the following self-consistent procedure:

1. Specify positions of the nuclei $\{\mathbf{R}_i\}$ and calculate $V(\mathbf{r})$
2. Construct an initial guess for $\rho(\mathbf{r})$, *e.g.* a superposition of atomic densities
3. Calculate $V_{\mathrm{H}}(\mathbf{r})$ and $V_{\mathrm{xc}}(\mathbf{r})$
4. Solve the KS equations numerically for $\{\lambda_i\}$ and $\{\phi_i(\mathbf{r})\}$
5. Construct the new $\rho(\mathbf{r})$ from $\{\phi_i(\mathbf{r})\}$

Steps 3-5 are repeated until the new and old electron densities are equal.

2.1.3. Local density approximation

In principle, solving the KS equations gives the exact ground state total energy provided that we know the exact form of the exchange-correlation functional for a polyatomic system, which we do not. We do, however, know its exact form for a simpler system, the homogeneous electron gas (HEG). (76–78) While the electron density in a polyatomic system rarely resembles a HEG, making this approximation for the

exchange-correlation functional in Eq. 2.13, *i.e.*

$$E_{\mathrm{xc}}^{\mathrm{LDA}}\left[\rho\left(\mathbf{r}\right)\right]=\int d\mathbf{r}\epsilon_{\mathrm{xc}}^{\mathrm{HEG}}\left[\rho\left(\mathbf{r}\right)\right]\rho\left(\mathbf{r}\right) \tag{2.20}$$

has proven quite successful. (79) Eq. 2.20 is known as the local density approximation (LDA). An intuitive reason for the success of the LDA is as follows. If you partition the electron density of a polyatomic system into infinitesimal volume elements, then, for each element, the density is nearly homogeneous. However, for systems with strongly spatially-varying electron densities, such as molecules, semiconductors, insulators, and crystals with defects, the LDA breaks down and more sophisticated exchange-correlation functionals are required.

2.1.4. Generalized gradient approximation

A natural extension of the LDA that favors density inhomogeneity is to include the gradient of the electron density in the exchange-correlation functional

$$E_{\mathrm{xc}}^{\mathrm{GGA}}\left[\rho\left(\mathbf{r}\right)\right]=\int d\mathbf{r}f\left[\rho\left(\mathbf{r}\right),\nabla\rho\left(\mathbf{r}\right)\right] \tag{2.21}$$

This is called the generalized gradient approximation (GGA) and it has been shown to improve total energies, enthalpies of atomization, and activation energies for chemical reactions. (80–86) The most widely used GGA is that of Perdew, Burke, and Ernzerhof (PBE). (87) They proposed a correlation energy of the form

$$E_{\mathrm{c}}^{\mathrm{GGA}}\left[\rho_{\uparrow},\rho_{\downarrow}\right]=\int d\mathbf{r}\rho\left[\epsilon_{\mathrm{c}}^{\mathrm{HEG}}\left(r_{s},\zeta\right)+H\left(r_{s},\zeta,t\right)\right] \tag{2.22}$$

where we have omitted the position dependence of the electron density for clarity, $\rho_{\uparrow}$ and $\rho_{\downarrow}$ are the up and down spin densities ($\rho=\rho_{\uparrow}+\rho_{\downarrow}$), $\epsilon_{\mathrm{c}}^{\mathrm{HEG}}$ is the correlation

energy per electron of a homogeneous electron gas, $r_s = (3/4\pi\rho)^{1/3}$ is the Wigner-Seitz radius, that is, the radius of a sphere whose volume is equal to the volume per electron, $\zeta = (\rho_\uparrow - \rho_\downarrow)/\rho$ is the relative spin polarization, t is a dimensionless density gradient, and H is a gradient contribution that satisfies three conditions:

1. If ρ is slowly varying, *i.e.* as $t \to 0$, then $H \propto t^2$. (88)

2. If ρ is fast varying, *i.e.* as $t \to \infty$, then $H \propto -\epsilon_{\mathrm{c}}^{\mathrm{HEG}}$. In other words, electron correlation vanishes.

3. If ρ increases, *i.e.* as $\rho(\mathbf{r}) \to \lambda^3 \rho(\lambda \mathbf{r})$ and $\lambda \to \infty$, then $H \propto \ln t^2$. (89)

For more details on the construction of H, the reader is referred to the original paper. (87)

On the other hand, they define the exchange energy as

$$E_{\mathrm{x}}^{\mathrm{GGA}} = \int d\mathbf{r} \rho \epsilon_{\mathrm{x}}^{\mathrm{HEG}}(\rho) F_{\mathrm{x}}(s) \tag{2.23}$$

where $\epsilon_{\mathrm{x}}^{\mathrm{HEG}} = -3e^2 k_{\mathrm{F}}/4\pi$ is the exchange energy per electron of a homogeneous electron gas, k_{F} is the radius of the Fermi surface, s is another dimensionless density gradient, and F_{x} is a gradient contribution that satisfies four additional conditions:

1. $E_{\mathrm{x}}[\rho_\uparrow, \rho_\downarrow] = (E_x[2\rho_\uparrow] + E_x[2\rho_\downarrow])/2$. (90)

2. If $\rho = \rho^{\mathrm{HEG}}$, *i.e.* if $s = 0$, then $F_{\mathrm{x}} = 1$.

3. If ρ is slowly varying, then the LDA is a better approximation for the exchange-correlation energy than the GGA. (91–93) Therefore, as $s \to 0$, $F_{\mathrm{x}} \propto 1 + \mu s^2$ where μ is a coefficient whose value is set so as to cancel the gradient contribution to correlation.

4. $E_{\mathrm{x}}\left[\rho_{\uparrow}, \rho_{\downarrow}\right] \geq E_{\mathrm{xc}}\left[\rho_{\uparrow}, \rho_{\downarrow}\right]$. (94)

This thesis comprises DFT calculations performed using the GGA of PBE.

2.1.5. Bloch's theorem and the plane wave basis set

In order to solve the KS equations numerically, we must choose a computational representation of the KS orbitals. This usually amounts to choosing a set of basis functions in which to expand the KS orbitals. A particularly fortuitous choice is a basis set consisting of plane waves, for reasons that will now be explained. In this thesis, we are primarily concerned with crystals and models of their surfaces that are periodic in three dimensions. When modeling a solid material with periodicity, it becomes unnecessary to consider the infinite number of atoms and electrons present in the macroscopic crystal. Instead, due to the underlying symmetry of the crystal, one can often identify a relatively small, fundamental repeating unit known as a primitive cell. Bloch's theorem states that we can use this periodicity to express the electronic wave function as

$$\phi_{n\mathbf{k}}\left(\mathbf{r}\right) = e^{i\mathbf{k}\cdot\mathbf{r}} u_{n\mathbf{k}}\left(\mathbf{r}\right) \tag{2.24}$$

where we have replaced i with n as the index of the energy band to avoid confusion with the imaginary number i, $\mathbf{k}$ is a wave vector, and $u_{n\mathbf{k}}\left(\mathbf{r}\right)$ is a function that has the same periodicity as the crystal, *i.e.* $u_{n\mathbf{k}}\left(\mathbf{r}\right) = u_{n\mathbf{k}}\left(\mathbf{r}+\mathbf{L}\right)$ where $\mathbf{L}$ is a translational vector of the primitive cell. (95) Plane waves are a natural choice to expand $u_{n\mathbf{k}}\left(\mathbf{r}\right)$ because they form a complete and orthonormal basis and can be made to satisfy the periodic boundary conditions of the crystal

$$f_{\mathbf{G}}\left(\mathbf{r}\right) = e^{i\mathbf{G}\cdot\mathbf{r}} \tag{2.25}$$

where $\mathbf{G}$ is a reciprocal lattice vector and $\mathbf{G} \cdot \mathbf{L} = 2\pi \times$ integer. Expanding $u_{n\mathbf{k}}(\mathbf{r})$ in plane waves,

$$u_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n\mathbf{k}}(\mathbf{G})\, e^{i\mathbf{G}\cdot\mathbf{r}} \tag{2.26}$$

where $c_{n\mathbf{k}}$ are coefficients that depend on $\mathbf{G}$, and inserting Eq. 2.26 into 2.24 gives

$$\phi_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_{n\mathbf{k}}(\mathbf{G})\, e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}} \tag{2.27}$$

Substitution of Eq. 2.27 into the KS equations (2.17) and integration over $\mathbf{r}$ produces

$$\sum_{\mathbf{G}'} \left[\frac{\hbar^2}{2m} |\mathbf{k}+\mathbf{G}|^2 \delta_{\mathbf{G}\mathbf{G}'} + V(\mathbf{G}-\mathbf{G}') + V_{\mathrm{H}}(\mathbf{G}-\mathbf{G}') + V_{\mathrm{xc}}(\mathbf{G}-\mathbf{G}') \right] c_{n\mathbf{k}}(\mathbf{G}') = \lambda_n c_{n\mathbf{k}}(\mathbf{G}) \tag{2.28}$$

where V, V_{H}, and V_{xc} are described in terms of their Fourier transforms. At this point, we will briefly discuss some technical aspects related to the sum over $\mathbf{G}'$ and the selection of $\mathbf{k}$. In principle, Eq. 2.28 contains an infinite sum over $\mathbf{G}'$ because there is an infinite number of repeats of the primitive cell in a macroscopic crystal. Of course, one can always choose to truncate this sum to make Eq. 2.28 computationally tractable. Typically, this is done in a systematic way by placing a cutoff on the kinetic energy of the plane waves included in the expansion, that is to say, the kinetic energy of a plane wave whose wave vector is $\mathbf{k}+\mathbf{G}$ must be less than some energy cutoff E_{cut}

$$E_{\mathbf{k}+\mathbf{G}} = \frac{\hbar^2}{2m} |\mathbf{k}+\mathbf{G}|^2 < E_{\mathrm{cut}} \tag{2.29}$$

where E_{cut} can be increased to achieve higher levels of total energy and force convergence. Truncation according to Eq. 2.29 works because higher kinetic energy plane waves generally correspond to unoccupied bands that do not contribute to the total

energy of the system. As for the selection of $\mathbf{k}$, the exact solution of the KS equations involves an integral over the Brillouin zone. However, like the infinite sum over $\mathbf{G}'$, this too is impossible. It has been shown that the integral over $\mathbf{k}$ can be replaced by a sum over a finite set of $\mathbf{k}$-points also known as a $\mathbf{k}$-point mesh. (96; 97) For practical DFT calculations, one usually benchmarks the convergence of the total energy and forces against the density of the $\mathbf{k}$-point mesh.

2.1.6. Pseudopotentials

Once the KS equations have been written in a plane wave basis, they can be solved by diagonalizing the matrix whose elements are given by Eq. 2.28. As the size of the system increases, however, so does the size of this matrix thus making the solution of the KS equations prohibitively expensive for all but the smallest systems containing elements with the fewest electrons. This scaling can be improved by recognizing the fact that the chemical bonding in a crystal is dominated by the valence electrons of the constituent atoms. (98–102) Another complication is that the wave functions of the valence electrons oscillate rapidly near the core, which increases the number of plane waves required for their description. To overcome these limitations, we can replace the hard ionic potential with a soft, effective potential that includes the interactions between the nucleus and core electrons and reproduces the eigenvalues and tails of the all-electron valence wave functions (see Fig. 2.1 for a P pseudopotential). This replacement is known as the pseudopotential approximation and it frequently accompanies the use of a plane wave basis. The first step of pseudopotential construction is to solve the radial KS equations for the all-electron (AE) eigenvalues $\{\lambda_l^{\mathrm{AE}}\}$ and wave functions $\{\phi_l^{\mathrm{AE}}(r)\}$ of a reference atomic electron configuration

$$\left(-\frac{d^2}{dr^2}+\frac{l(l+1)}{r^2}-\frac{2Z}{r}+V_{\mathrm{Hxc}}[\rho(r)]\right)\phi_l^{\mathrm{AE}}(r)=\lambda_l\phi_l^{\mathrm{AE}}(r) \tag{2.30}$$

where l is the orbital angular momentum quantum number and $V_{\mathrm{Hxc}} = V_{\mathrm{H}} + V_{\mathrm{xc}}$. Next, pseudo-wave functions (PS) are designed such that the following five rules are upheld:

1. $\{\lambda_l^{\mathrm{AE}}\}_{\mathrm{ref}} = \{\lambda_l^{\mathrm{PS}}\}_{\mathrm{ref}}$.
2. $\phi_{l,\mathrm{ref}}^{\mathrm{AE}} = \phi_{l,\mathrm{ref}}^{\mathrm{PS}}$ for $r \geq r_{\mathrm{c}}$ where r_{c} is a user-defined cutoff radius.
3. $N_{l,\mathrm{ref}}^{\mathrm{AE}} = N_{l,\mathrm{ref}}^{\mathrm{PS}}$ where $N_l = \int_0^{r_{\mathrm{c}}} dr r^2 \left|\phi_l\left(r\right)\right|^2$.
4. $D_l^{\mathrm{AE}}\left(\lambda, r_{\mathrm{c}}\right) = D_l^{\mathrm{PS}}\left(\lambda, r_{\mathrm{c}}\right)$ where $D_l\left(\lambda, r_{\mathrm{c}}\right) = r\frac{d}{dr}\ln\phi_l\left(\lambda, r_{\mathrm{c}}\right)$.
5. $\frac{\partial}{\partial\lambda}D_l^{\mathrm{AE}}\left(\lambda, r_{\mathrm{c}}\right) = \frac{\partial}{\partial\lambda}D_l^{\mathrm{PS}}\left(\lambda, r_{\mathrm{c}}\right)$.

If a pseudo-wave function obeys rules 3 and 5, then the resulting pseudopotentials are called "norm-conserving". (103) Once the pseudo-wave functions have been designed, screened (scr) semi-local pseudopotentials are calculated by inverting Eq. 2.30, which yields

$$V_l^{\mathrm{scr}}\left(r\right) = \lambda_l - \frac{l\left(l+1\right)}{r^2} + \frac{1}{\phi_l^{\mathrm{PS}}\left(r\right)}\frac{d^2}{dr^2}\phi_l^{\mathrm{PS}}\left(r\right) \tag{2.31}$$

The screened pseudopotentials are then descreened to give the ionic pseudopotentials

$$V_l^{\mathrm{PS}}\left(r\right) = V_l^{\mathrm{scr}}\left(r\right) - V_{\mathrm{Hxc}}\left[\rho^{\mathrm{val}}\left(r\right)\right] \tag{2.32}$$

where $\rho^{\mathrm{val}}\left(r\right)$ is the valence electron density. The ionic pseudopotentials of P are shown at the bottom of Fig. 2.1. Eq. 2.32 suggests that each orbital angular momentum of the wave function experiences a different ionic potential.

It turns out to be more efficient to transform semi-local pseudopotentials into a

fully separable, nonlocal (NL) form

$$\hat{V}_{\mathrm{NL}} = V_{\mathrm{loc}}(r) + \sum_{l} \frac{\left|\Delta V_l(r)\,\phi_l^{\mathrm{ref}}\right\rangle \left\langle \phi_l^{\mathrm{ref}} \Delta V_l(r)\right|}{\left\langle \phi_l^{\mathrm{ref}}\left|\Delta V_l(r)\right|\phi_l^{\mathrm{ref}}\right\rangle} \tag{2.33}$$

where $V_{\mathrm{loc}}(r)$ is an arbitrarily chosen local potential and

$$\Delta V_l(r) = V_l^{\mathrm{PS}}(r) - V_{\mathrm{loc}}(r) \tag{2.34}$$

This is the well-known Kleinman-Bylander nonlocal form. (104) Further, it has been shown that the transferability of a nonlocal pseudopotential, *i.e.* its ability to reproduce all-electron calculations of non-reference states, can be enhanced by adding an augmentation function $A(r)$ to the local potential

$$\hat{V}_{\mathrm{DNL}} = V_{\mathrm{loc}}(r) + A(r) + \sum_{l} \frac{\left|\left(\Delta V_l(r) - A(r)\right)\phi_l^{\mathrm{ref}}\right\rangle \left\langle \phi_l^{\mathrm{ref}}\left(\Delta V_l(r) - A(r)\right)\right|}{\left\langle \phi_l^{\mathrm{ref}}\left|\left(\Delta V_l(r) - A(r)\right)\right|\phi_l^{\mathrm{ref}}\right\rangle} \tag{2.35}$$

where DNL stands for designed nonlocal. (105) It is straightforward to show that Eq. 2.35 does not affect the reference state and enables tunable enhancement of transferability via the augmentation function. The calculations presented herein use plane waves and designed nonlocal (105) optimized pseudopotentials. (106)

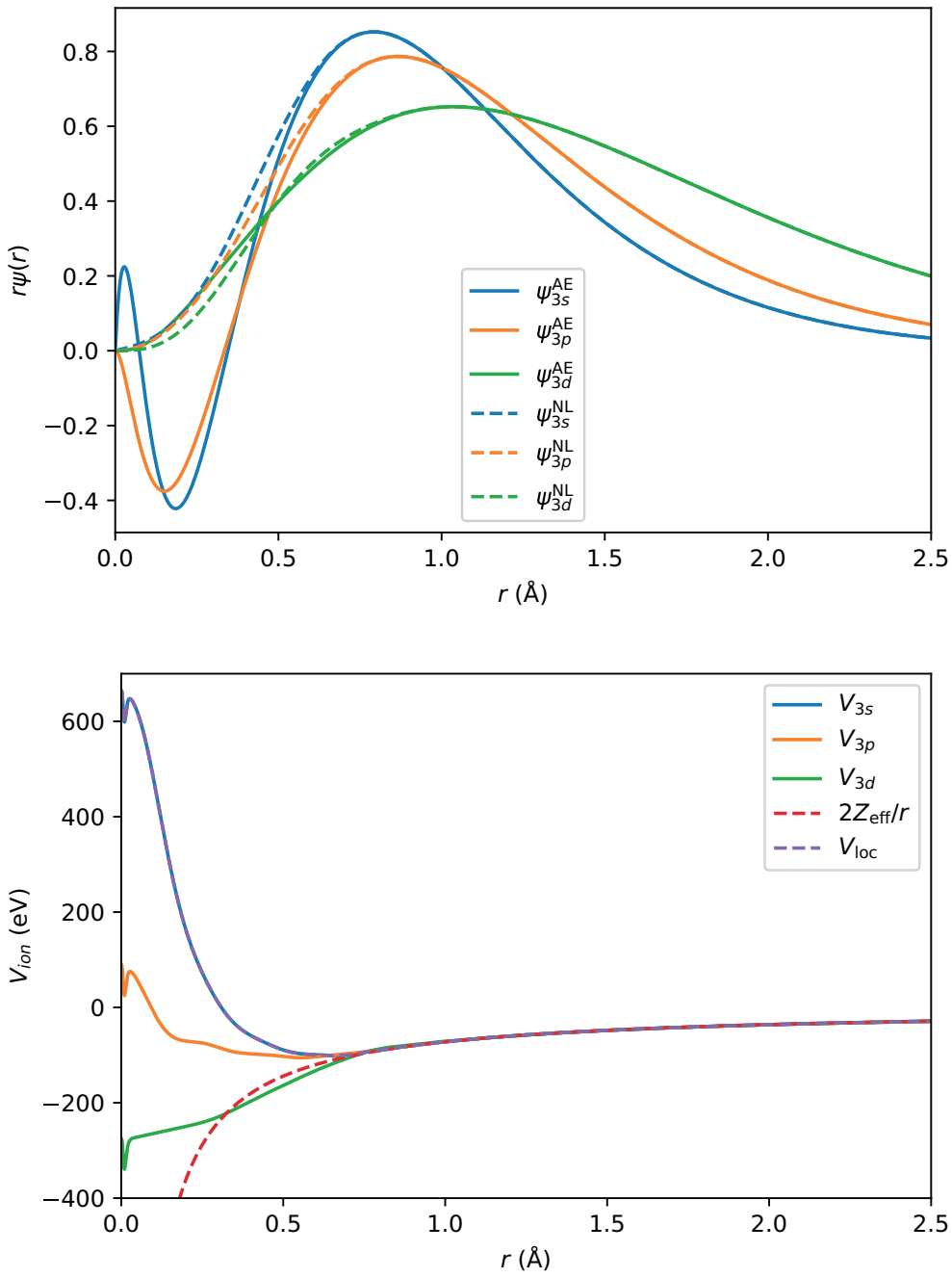

Figure 2.1: Pseudopotential of P. (Top) All-electron (AE) wave functions and non-local (NL) pseudo-wave functions of the P $3s$, $3p$, and $3d$ states. NL pseudo-wave functions are nodeless and smoother than the AE wave functions. (Bottom) Ionic pseudopotentials V_{3l}, true atomic potential (shaded red), and local potential V_{loc}. Ionic pseudopotentials do not diverge as $r \to 0$ and are smoother than the true atomic potential.

2.1.7. van der Waals interactions

A drawback of the LDA and GGA is that the resulting exchange-correlation functionals are unable to correctly describe the nonlocal electron correlations responsible for van der Waals (vdW) interactions. (107; 108) For example, GGAs tend to underestimate their strength. (109; 110) This is of great concern because vdW interactions can influence the thermodynamics of chemical reactions, which is the central topic of this thesis. (111–119) In order to rectify the GGA's treatment of vdW interactions, many different corrections have been proposed. (120–128) Of those, perhaps the simplest is the semiempirical correction of the form

$$E_{\mathrm{vdW}} = -\frac{1}{2}\sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{\mathbf{L}}\frac{C_{6ij}}{r_{ij,\mathbf{L}}^{6}}f_{d,6}\left(r_{ij,\mathbf{L}}\right) \tag{2.36}$$

where C_{6ij} are the vdW coefficients of the atom pair ij, $r_{ij,\mathbf{L}}$ is the distance between atom i in the primitive cell and atom j in a neighboring cell translated from the origin by $\mathbf{L}$, and $f_{d,6}$ is a damping function that avoids the counting of vdW interactions between atoms at chemical bonding distances. (129) The vdW coefficients of atom pairs can be written as the geometric mean of atomic vdW coefficients

$$C_{6ij} = \sqrt{C_6^i C_6^j} \tag{2.37}$$

where C_6 of atom i is given by

$$C_6^i = 0.05 \times m I_p^i \alpha_p^i \tag{2.38}$$

Here, m = 2, 10, 18, and 36 for rows 1, 2, 3, and 4 of the periodic table, I_p is the atomic ionization potential, and α_p is the static dipole polarizability. Eq. 2.38 has a

quantum mechanical foundation as it is derived from the London formula. (130–132) The form of the damping function is flexible, however, it is typically expressed as

$$f_{d,6}\left(r_{ij}\right) = \frac{s_6}{1 + e^{-d(r_{ij}/R_{0ij}-1)}} \tag{2.39}$$

where s_6 is a scaling factor that depends on the exchange-correlation functional, d is a damping parameter, and R_{0ij} is

$$R_{0ij} = R_0^i + R_0^j \tag{2.40}$$

where R_0^i is the vdW radius of atom i. (129) To apply this vdW correction, Eq. 2.36 is simply added to the KS total energy. The forces, on the other hand, are computed from finite differences.

2.2. *Ab initio* thermodynamics

The primary goal of this thesis is to determine the thermodynamic barriers associated with surface chemical reactions, specifically those occurring in an electrochemical environment. Such an endeavor requires knowledge of the Gibbs free energies of the reactants, products, and intermediate species in the chemical reaction of interest. The Gibbs free energy G is the Legendre transform of the internal energy $U\left(S, V, N\right)$, which results in

$$G = U - TS + pV \tag{2.41}$$

$$dG = -SdT + Vdp + \mu dN \tag{2.42}$$

where the natural variables of G are the temperature T, pressure p, and the number of particles N, S is the entropy, V is the volume, μ is the chemical potential, and dG

is the differential form of G for a quasistatic process. $G(T, p, N)$ is an example of a thermodynamic potential, which is a scalar that represents the state of a thermodynamic system. It describes the capacity of a system to do non-mechanical work and is minimized, in other words, $dG = 0$, when the system reaches chemical equilibrium at constant T and p. It follows from this that G must decrease for a chemical reaction to be spontaneous.

For reasons that will soon be evident, it is important to know the relationship between μ, T, and p. μ is defined as

$$\mu = \left(\frac{\partial G}{\partial N}\right)_{T,p} \tag{2.43}$$

Multiplying both sides of Eq. 2.43 by dN and integrating, we find that G and dG can be expressed as

$$G = \mu N = \sum_{i=1}^{I} \mu_i N_i \tag{2.44}$$

$$dG = \sum_{i=1}^{I} \mu_i dN_i + \sum_{i=1}^{I} N_i d\mu_i \tag{2.45}$$

where the sum is over the types of particles, of which there are I. By setting Eqs. 2.42 and 2.45 equal, we obtain

$$\sum_{i=1}^{I} N_i d\mu_i = -SdT + Vdp \tag{2.46}$$

which is known as the Gibbs-Duhem equation. Eq. 2.46 is useful because it allows us to calculate the change in μ upon isothermal compression/expansion and isobaric heating/cooling.

First, let us consider the isothermal compression/expansion of an ideal gas. This is relevant to our study of surface chemical reactions because the reactants/products may be in the gas phase and, under certain circumstances, behave like ideal gases. For constant T, Eq. 2.46 becomes

$$Nd\mu = Vdp \tag{2.47}$$

for a single phase ideal gas. Replacing V in Eq. 2.47 with the ideal gas equation of state and integrating yields

$$\Delta\mu = k_{\mathrm{B}}T \ln \frac{p_f}{p_i} \tag{2.48}$$

where k_{B} is the Boltzmann constant, and p_f and p_i are the final and initial pressures, respectively.

Since we are interested in the chemical reactivity of crystals, we should also consider the isothermal compression/expansion of solids. Unfortunately, solids do not have a simple equation of state like ideal gases. Progress can be made in deriving $\Delta\mu$, however, by noting that the isothermal compressibility of solids is small, *i.e.*

$$\beta_T = -\frac{1}{V}\left(\frac{\partial V}{\partial p}\right)_T \approx 0 \tag{2.49}$$

From Eq. 2.49, it is clear that the volume is constant with respect to quasistatic changes in pressure. Using this result, we integrate Eq. 2.47

$$\Delta\mu = \frac{1}{N}\int Vdp = \frac{1}{\rho}\left(p_f - p_i\right) \tag{2.50}$$

where ρ is the number density of particles. For example, if the pressure on Si is increased by 1 bar, then $\Delta\mu = 1.3 \times 10^{-5}$ eV/atom at 298 K. This is small compared

to the free energy changes of chemical reactions and therefore can be safely neglected.

Now, let us consider the isobaric heating/cooling of a system. For constant p, the integrated form of Eq. 2.46 is

$$\Delta\mu = -\frac{1}{N}\int SdT \tag{2.51}$$

We do not know, however, an equation of state that relates S and T. Another way to approach this problem is to start with the thermodynamic potential that corresponds to constant p, *i.e.* the enthalpy

$$H = U + pV \tag{2.52}$$

$$dH = TdS + Vdp + \mu dN \tag{2.53}$$

For an isobaric process,

$$\Delta H = \int TdS = \int c_p dT \tag{2.54}$$

where we have replaced TdS with the definition of the isobaric heat capacity

$$c_p = T\left(\frac{\partial S}{\partial T}\right)_p \tag{2.55}$$

ΔH in Eq. 2.54 is called the integrated heat capacity (IHC). Combining Eqs. 2.41, 2.44, 2.52, and 2.54, one obtains

$$\Delta\mu = \frac{1}{N}\left(\int c_p dT - \Delta\left(TS\right)\right) \tag{2.56}$$

Experimental values of the IHC and $\Delta\left(TS\right)$ for different compounds and temperatures are published in thermochemical tables. (133; 134)

We are now ready to calculate the free energy changes of chemical reactions. Consider the general chemical reaction

$$\sum_{i=1}^{N_{\mathrm{R}}} m_i \mathrm{R}_i \rightleftharpoons \sum_{j=1}^{N_{\mathrm{P}}} n_j \mathrm{P}_j \tag{2.57}$$

where N_{R} and N_{P} are the number of reactants (R) and products (P), i and j are their indices, and m and n are their stoichiometric coefficients, respectively. The free energy change of this reaction (rxn) is

$$\begin{aligned}
\Delta G_{\mathrm{rxn}} &= \sum_{j=1}^{N_{\mathrm{P}}} n_j G_j - \sum_{i=1}^{N_{\mathrm{R}}} m_i G_i \\
&= \sum_{j=1}^{N_{\mathrm{P}}} n_j H_j - \sum_{j=1}^{N_{\mathrm{P}}} n_j T S_j + \sum_{j=1}^{N_{\mathrm{P}}} n_j \mathrm{IHC}_j\left(T\right) \\
&- \sum_{i=1}^{N_{\mathrm{R}}} m_i H_i + \sum_{i=1}^{N_{\mathrm{R}}} m_i T S_i - \sum_{i=1}^{N_{\mathrm{R}}} m_i \mathrm{IHC}_i\left(T\right) \\
&= \Delta H_{\mathrm{rxn}} - T\Delta S_{\mathrm{rxn}} + \Delta \mathrm{IHC}_{\mathrm{rxn}}\left(T\right)
\end{aligned} \tag{2.58}$$

where ΔH_{rxn} is the enthalpy change and $\Delta \mathrm{IHC}_{\mathrm{rxn}}$ is the integrated heat capacity change. Practically speaking, H is approximated by the DFT total energy, which corresponds to 0 K. S and the IHC are taken from thermochemical tables. (133; 134)

For surface chemical reactions, the free energy changes associated with the vibrations of surface atoms and adsorbates must also be taken into account. To understand why this is necessary, consider the following example. In the gas phase, H_2 has a stretching frequency of 4401 cm^{-1}. (135) On the Pt(111) surface, however, H has two vibrational frequencies of 550 cm^{-1} and 1230 cm^{-1}, which correspond to asymmetric and symmetric Pt-H stretching modes, respectively. (136) This reduction in the vibrational frequencies of H leads to a 0.01 eV increase in ΔG_{rxn}, which is

small but nonzero and therefore important for obtaining quantitative agreement with experiments. We start from the partition function Z of a harmonic solid

$$Z = \prod_{i=1}^{3N} \frac{e^{-\hbar\omega_i/2k_\mathrm{B}T}}{1 - e^{-\hbar\omega_i/2k_\mathrm{B}T}} \tag{2.59}$$

where ω_i is the vibrational frequency of oscillator i. Z is related to the vibrational free energy via

$$F_\mathrm{vib} = -k_\mathrm{B}T \ln Z = \frac{1}{2}\sum_{i=1}^{3N} \hbar\omega_i + k_\mathrm{B}T \sum_{i=1}^{3N} \ln\left(1 - e^{-\hbar\omega_i/k_\mathrm{B}T}\right) \tag{2.60}$$

where the first sum on the right-hand side is the zero-point energy (ZPE). At high temperatures, the harmonic approximation breaks down because the atomic vibrations can now sample regions of the potential energy surface where anharmonic effects are important. Therefore, one must be careful, when using Eq. 2.60, to ensure that the system is in the harmonic regime at the temperatures of interest. Here, ω_i were calculated using density-functional perturbation theory (DFPT). (137)

2.2.1. Bulk phase diagrams

Before we can study chemical reactions on a surface, we first have to determine the conditions under which the bulk crystal is stable. This is done by constructing a bulk phase diagram, which plots the thermodynamically preferred bulk phase as a function of the chemical potentials of the constituent elements. Take, for example, the hypothetical binary compound $\mathrm{A}_x\mathrm{B}_y$, which has three stable chemical compositions, $x_1/y_1 < x_2/y_2 < x_3/y_3$. The chemical equilibrium of $\mathrm{A}_{x_2}\mathrm{B}_{y_2}$ is

$$x_2\mathrm{A} + y_2\mathrm{B} \rightleftharpoons A_{x_2}B_{y_2} \tag{2.61}$$

The free energy of $A_{x_2}B_{y_2}$ formation can be written as

$$\Delta G_{f,A_{x_2}B_{y_2}} = x_2\Delta\mu_A + y_2\Delta\mu_B \tag{2.62}$$

Here, $\Delta\mu = \mu - E_{\text{std}}$ where E_{std} is the DFT total energy of the standard state. Eq. 2.62 shows that the equilibrium in Eq. 2.61 is a line in the ($\Delta\mu_A$, $\Delta\mu_B$) plane. If one phase occurs on a line, then two phases coexist at the point where their lines intersect. The coordinates of the intersection of $A_{x_1}B_{y_1}$ and $A_{x_2}B_{y_2}$ is calculated by solving the following matrix equation

$$\begin{bmatrix} x_1 & y_1 \\ x_2 & y_2 \end{bmatrix} \begin{bmatrix} \Delta\mu_A \\ \Delta\mu_B \end{bmatrix} = \begin{bmatrix} \Delta G_{f,A_{x_1}B_{y_2}} \\ \Delta G_{f,A_{x_2}B_{y_2}} \end{bmatrix} \tag{2.63}$$

$$\mathbf{C}\vec{\Delta\mu} = \Delta\vec{G}_f \tag{2.64}$$

$$\vec{\Delta\mu} = \mathbf{C}^{-1}\Delta\vec{G}_f \tag{2.65}$$

The same procedure is repeated for the equilibrium between $A_{x_2}B_{y_2}$ and $A_{x_3}B_{y_3}$. The resulting sets of chemical potentials, $\vec{\Delta\mu}_{12}$ and $\vec{\Delta\mu}_{23}$, enclose the region where $A_{x_2}B_{y_2}$ is the preferred phase. Fig. 2.2 shows a schematic bulk phase diagram for A_xB_y.

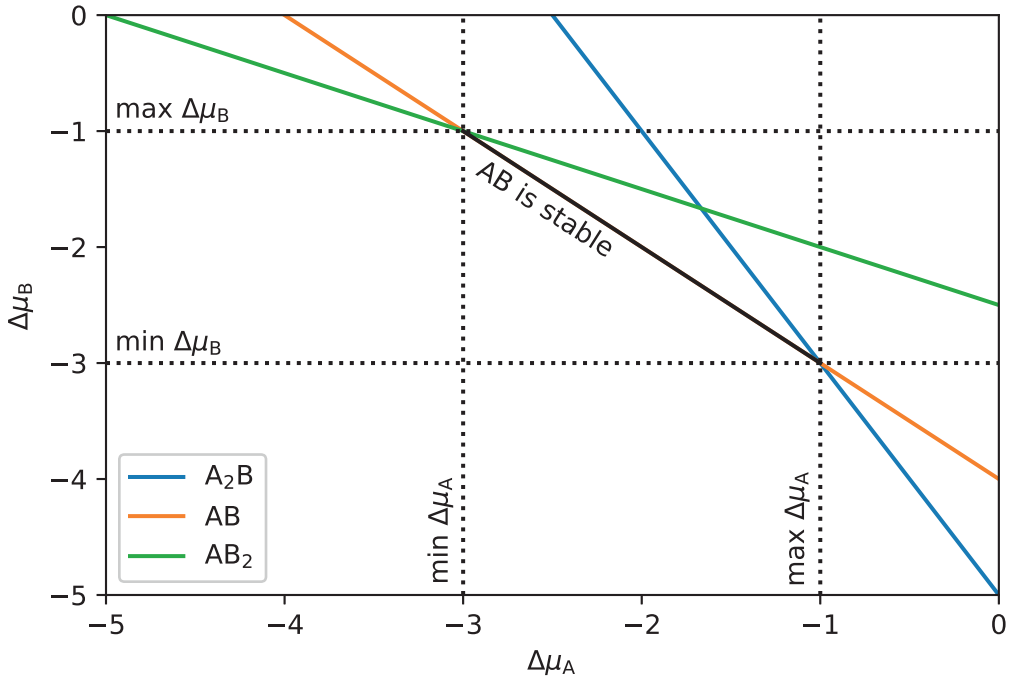

Figure 2.2: Bulk phase diagram of a hypothetical binary compound A_xB_y with three stable compositions A_2B, AB, and AB_2. Dotted lines enclose the region of $\Delta\mu_A$ and $\Delta\mu_B$ where AB is stable.

2.2.2. Surface phase diagrams

Now that we have established the regions of chemical potential where the bulk is stable, we want to determine the surfaces that are stable therein. To do this, we first have to construct a computational model of a surface. In the context of DFT calculations that are periodic in three dimensions, it is customary to employ a slab model like the one shown in Fig. 2.3. This slab has three regions: bulk, surface, and vacuum. The positions of the atoms in the bottom few layers of the slab are fixed to replicate a stable, underlying bulk. Additionally, the height of the vacuum must be large enough so that the top and bottom layers of the slab do not interact. The free energy of the slab can then be expressed as

$$G_{\text{slab}} = G_{\text{bulk}} + A\gamma \tag{2.66}$$

where G_{bulk} is the free energy of the equivalent amount of bulk, A is the surface area, and γ is the surface energy per unit area. Inserting Eq. 2.44 and solving for γ gives

$$\gamma = \frac{1}{A}\left[G_{\text{slab}} - \sum_{i=1}^{I} N_i \mu_i\right] \tag{2.67}$$

It has been shown that G_{slab} can be safely replaced by E_{slab}, which is the DFT total energy of the slab. (57) It is also convenient to use the relative chemical potentials introduced in Eq. 2.62. Upon making these substitutions, Eq. 2.67 becomes

$$\gamma = \frac{1}{A}\left[E_{\text{slab}} - \sum_{i=1}^{I} N_i \left(\Delta\mu_i + E_{i,\text{std}}\right)\right] \tag{2.68}$$

The surface energy of our hypothetical binary compound A_xB_y is therefore

$$\gamma = \frac{1}{A}\left[E_{\text{slab}} - N_{\text{A}}\Delta\mu_{\text{A}} - N_{\text{A}}E_{\text{A,std}} - N_{\text{B}}\Delta\mu_{\text{B}} - N_{\text{B}}E_{\text{B,std}}\right] \tag{2.69}$$

Since the surface is in equilibrium with bulk A_xB_y and the bulk is in equilibrium with A and B, we need only specify one chemical potential as the two are related by Eq. 2.62. In other words, we can replace $\Delta\mu_{\text{A}}$ with

$$\Delta\mu_{\text{A}} = \frac{\Delta G_{f,\text{A}_x\text{B}_y}}{x} - \frac{y}{x}\Delta\mu_{\text{B}} \tag{2.70}$$

Approximating ΔG_f with the DFT total energy of formation (ΔE_f), we can write Eq. 2.69 as

$$\gamma = \gamma_0 + \frac{\Gamma_{\text{B}}}{A}\Delta\mu_{\text{B}} \tag{2.71}$$

where

$$\gamma_0 = \frac{1}{A}\left[E_{\text{slab}} - \frac{N_{\text{A}}E_{\text{A}_x\text{B}_y}}{x} + \Gamma_{\text{B}}E_{\text{B,std}}\right] \tag{2.72}$$

$$\Gamma_{\text{B}} = \frac{N_{\text{A}}y}{x} - N_{\text{B}} \tag{2.73}$$

Using Eqs. 2.71 through 2.73, we can now calculate surface energies directly from first principles. The remaining challenge is to generate a database of slabs that efficiently samples surface phase space. This can be done in one of two ways. The first is to build slabs using the chemical intuition of Pauling's rules, (99; 138) crystal (139) and ligand field theory, (140) Tasker's criteria for the electrostatic energy convergence of ionic crystal surfaces, (141) *etc.* The second is to use crystal structure prediction software based on evolution algorithms, (59–61) particle swarm optimization, (63; 64) or *ab initio* grand canonical Monte Carlo. (68) The latter will be discussed in section

2.3.

Fig. 2.3 shows a schematic surface phase diagram for A_xB_y that consists of three phases, each with a different composition (N_A/N_B) relative to that of the bulk (x/y). The dotted lines enclose the region where bulk A_xB_y is stable. In this region, there are two stable slabs (shaded orange and blue) making up what is called the surface energy convex hull. The first (shaded orange) has the same composition as the bulk and therefore its surface energy does not depend on $\Delta\mu_B$ (because $\Gamma_B = 0$). The second (shaded blue) is enriched with B, which causes its surface energy to decrease with increasing $\Delta\mu_B$ (because $\Gamma_B < 0$). This means that the B-enriched slab becomes more stable as B becomes more available.

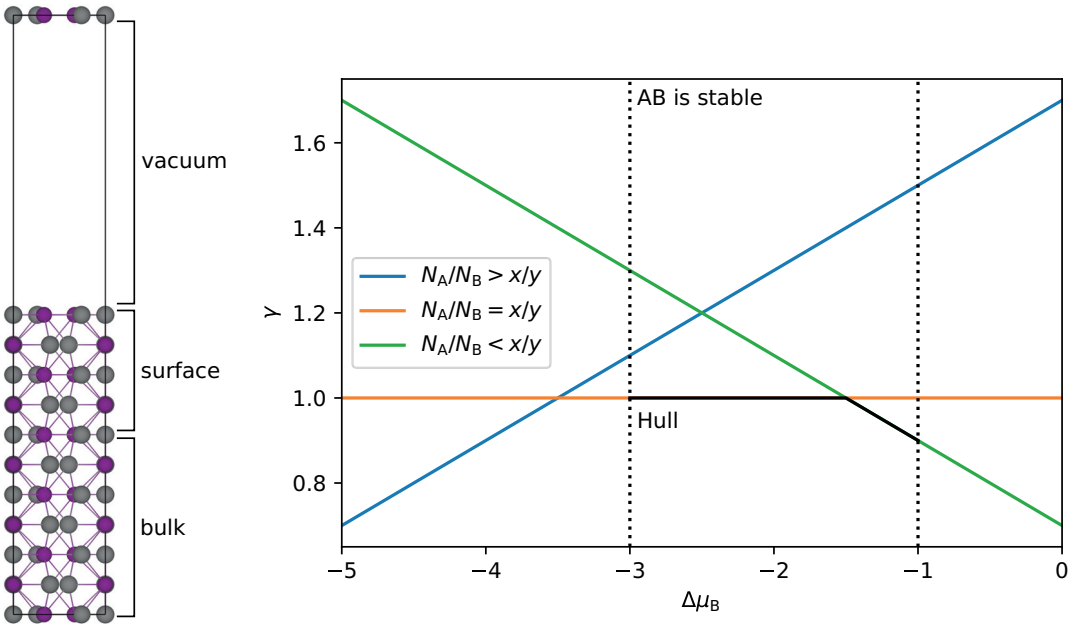

Figure 2.3: Slab model (left) and surface phase diagram (right) of a hypothetical binary compound A_xB_y. (Left) Silver and purple spheres correspond to A and B, respectively. (Right) Surface energy *vs.* the relative chemical potential of B. Three surface compositions are shown, two of which (orange and green) are stable in the region of $\Delta\mu_B$ where bulk AB is stable.

2.2.3. Pourbaix diagrams

The last two sections deal with the case when A_xB_y is in equilibrium with $A(s)$ and $B(s)$. This situation is relevant only if A_xB_y is synthesized, characterized, and evaluated in the solid-state. If it is to be synthesized solvothermally and/or used as an electrocatalyst, however, one must instead consider the equilibrium between A_xB_y, $A(aq)$, and $B(aq)$. Under these circumstances, Eq. 2.61 becomes

$$x\mathrm{A}\,(s) + y\mathrm{B}\,(s) \rightleftharpoons x\mathrm{A}\,(aq) + y\mathrm{B}\,(aq) \rightleftharpoons \mathrm{A}_x\mathrm{B}_y \tag{2.74}$$

where the first equilibrium corresponds to the aqueous solvation of A and B. Taking one component at a time, the aqueous solvation of A can be written as

$$\mathrm{A}\,(s) + n_{\mathrm{w}}\mathrm{H_2O} \rightleftharpoons [\mathrm{H}_x\mathrm{AO}_y]^z + n_{\mathrm{H}}\mathrm{H}^+ + n_e e^- \tag{2.75}$$

where n_{w} is the number of water molecules, $[\mathrm{H}_x\mathrm{AO}_y]^z$ is the aqueous phase of A, z is its charge, n_{H} is the number of protons, and n_e is the number of electrons. This is a generic redox process where A goes from a neutral charge state to $q_{\mathrm{A}} = -x + 2y + z$. Given the composition (x and y) and charge (z) of this aqueous phase, we can calculate n_{w}, n_{H}, and n_e using the following transformation

$$\begin{bmatrix} 0 & 1 & 0 \\ -1 & 2 & 0 \\ -1 & 2 & 1 \end{bmatrix} \begin{bmatrix} x \\ y \\ z \end{bmatrix} = \begin{bmatrix} n_{\mathrm{w}} \\ n_{\mathrm{H}} \\ n_e \end{bmatrix} \tag{2.76}$$

The free energy change of Eq. 2.75 is

$$\Delta G_{\mathrm{A,solv}} = \Delta G^{\circ}_{\mathrm{A,solv}} + k_{\mathrm{B}}T \ln a_{[\mathrm{H}_x\mathrm{AO}_y]^z} - 2.303 n_{\mathrm{H}} k_{\mathrm{B}} T p\mathrm{H} - n_e U \tag{2.77}$$

where a is the activity, U is the potential relative to that of the standard hydrogen electrode (SHE), and

$$\Delta G^{\circ}_{\mathrm{A,solv}} = G^{\circ}_{[\mathrm{H}_x\mathrm{AO}_y]^z} + n_{\mathrm{H}} G^{\circ}_{\mathrm{H}} + n_e G^{\circ}_e - G^{\circ}_{\mathrm{A}} - n_{\mathrm{w}} G^{\circ}_{\mathrm{H_2O}} \tag{2.78}$$

where G° is the free energy at $a_{[\mathrm{H}_x\mathrm{AO}_y]^z} = 1$ M, $p\mathrm{H} = 0$, and $U = 0$ V. In practice, $\Delta G^{\circ}_{\mathrm{solv}}$ is taken from experimental thermochemical tables because it is difficult to calculate the free energy of aqueous species. (1)

There may be several possible aqueous phases, each with a different oxidation state of A. For example, if A is Ni, then $[\mathrm{H}_x\mathrm{AO}_y]^z$ can be $\mathrm{Ni}(s)$, $\mathrm{Ni}^{2+}(aq)$, $\mathrm{NiO}(s)$, $\mathrm{NiOH}^{+}(aq)$, $\mathrm{Ni(OH)_2}(s)$, and $\mathrm{Ni(OH)_2}(aq)$. (1) The most stable aqueous phase of A is plotted as a function of U and $p\mathrm{H}$ in Fig. 2.4. This is called a Pourbaix diagram. (142) In general, A is oxidized as U and $p\mathrm{H}$ are increased. Using Eqs. 2.62 and 2.75, the free energy of forming $\mathrm{A}_x\mathrm{B}_y$ in aqueous solution is

$$\Delta G_{f,\mathrm{A}_x\mathrm{B}_y} = x \min\left(\Delta G_{\mathrm{A,solv}}\right) + y \min\left(\Delta G_{\mathrm{B,solv}}\right) \tag{2.79}$$

where the chemical potentials of A and B have been replaced by the solvation free energies of their most stable phases, *i.e.* those that minimize ΔG_{solv}. This equilibrium condition forms a closed region in the Pourbaix diagram wherein $\mathrm{A}_x\mathrm{B}_y$ is stable with respect to aqueous solvation of A and B (see dotted lines in Fig. 2.4).

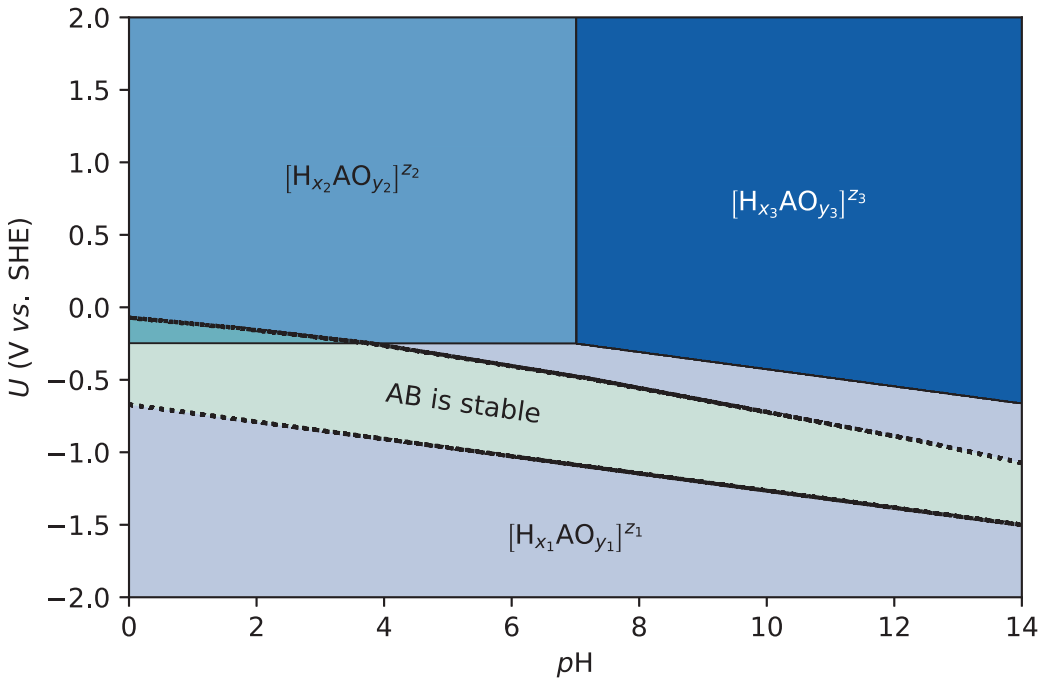

Figure 2.4: Pourbaix diagram of a hypothetical element A with three stable aqueous phases. Dotted lines enclose the region of U and pH where AB is stable.

2.2.4. Aqueous surface phase diagrams

In the preceding section, we showed how to calculate the region of U and pH where bulk A_xB_y is stable against aqueous solvation. Before we proceed, however, it is prudent to reiterate that bulk stability is a precondition for surface stability. At this time, we will extend our treatment of aqueous equilibria to the surfaces of A_xB_y. There are many surface chemical reactions that can occur in aqueous solution. These include (1) the dissolution and deposition of A and B, (2) the adsorption and desorption of H_aO_b species (H, OH, O, and H_2O), and (3) the combinations thereof (see Fig. 2.5). For brevity, we will focus on scenarios 1 and 2, that being said, the following approaches are transferable to scenario 3, *mutatis mutandis*.

Starting with scenario 1, the equilibrium between a surface A atom and aqueous solution is

$$\mathrm{SA} + n_{\mathrm{w}}\mathrm{H_2O} \rightleftharpoons \mathrm{S} + [\mathrm{H}_x\mathrm{AO}_y]^z + n_{\mathrm{H}}\mathrm{H}^+ + n_e e^- \tag{2.80}$$

where SA is the surface plus a surface A atom. The forward and backward reactions here are the dissolution and deposition of A, respectively. Note that it is trivial to rewrite Eq. 2.80 in terms of B. It is convenient to regard surface A dissolution as two consecutive half-reactions, *i.e.*

$$\mathrm{SA} \rightleftharpoons \mathrm{S} + \mathrm{A}\,(s) \tag{2.81}$$

and the forward reaction in Eq. 2.75. The first reaction (Eq. 2.81) involves the desorption of a surface A atom and its subsequent placement in a reservoir of A(s). It has been shown that the free energy of surface atom desorption (des) can be replaced by the DFT total energy (ΔE_{des}) because the entropy changes for solid-state chemical reactions are small. (57) Based on this simplification, the free energy change upon

dissolution (diss) of a surface A atom is

$$\Delta G_{\mathrm{A,diss}}\left(U, p\mathrm{H}\right) = \Delta E_{\mathrm{A,des}} + \Delta G_{\mathrm{A,solv}}\left(U, p\mathrm{H}\right) \tag{2.82}$$

where similar expressions exist for the deposition of A, and the dissolution and deposition of B. For deposition, however, the first and second terms in Eq. 2.82 correspond to the DFT total energy of adsorption and the free energy of precipitating a neutral atom, respectively.

Moving on to scenario 2, the chemical equation for $\mathrm{H}_a\mathrm{O}_b$ adsorption is

$$\mathrm{S} + b\mathrm{H_2O} \rightleftharpoons \mathrm{SO}_b\mathrm{H}_a + (2b - a)\left(\mathrm{H}^+ + e^-\right) \tag{2.83}$$

where $\mathrm{SO}_b\mathrm{H}_a$ is the surface plus an adsorbed $\mathrm{H}_a\mathrm{O}_b$ species. Clearly, the reverse reaction is the desorption of $\mathrm{H}_a\mathrm{O}_b$. The free energy of $\mathrm{H}_a\mathrm{O}_b$ adsorption is given by

$$\begin{aligned}\Delta G_{\mathrm{ads}}\left(U, p\mathrm{H}\right) = \Delta E_{\mathrm{ads}} + \Delta\mathrm{ZPE}_{\mathrm{ads}} - T\Delta S_{\mathrm{ads}} + \Delta\mathrm{IHC}_{\mathrm{ads}}\left(T\right) \\ - \alpha\left(2.303 k_B T p\mathrm{H} + U\right) - b k_{\mathrm{B}} T \ln p_{\mathrm{H_2O}(g)}\end{aligned} \tag{2.84}$$

where

$$\alpha = 2b - a \tag{2.85}$$

$$\Delta E_{\mathrm{ads}} = E_{\mathrm{SO}_b\mathrm{H}_a} + \alpha E_{\mathrm{H_2}}/2 - E_{\mathrm{S}} - bE_{\mathrm{H_2O}} \tag{2.86}$$

$$\Delta\mathrm{ZPE}_{\mathrm{ads}} = \mathrm{ZPE}_{\mathrm{SO}_b\mathrm{H}_a} + \alpha\mathrm{ZPE}_{\mathrm{H_2}(g)}/2 - \mathrm{ZPE}_{\mathrm{S}} - b\mathrm{ZPE}_{\mathrm{H_2O}(g)} \tag{2.87}$$

$$\Delta S_{\mathrm{ads}}\left(T\right) \approx \alpha S^{\circ}_{\mathrm{H_2}(g)}\left(T\right)/2 - bS^{\circ}_{\mathrm{H_2O}(g)}\left(T\right) \tag{2.88}$$

$$\Delta\mathrm{IHC}_{\mathrm{ads}}\left(T\right) \approx \alpha\mathrm{IHC}^{\circ}_{\mathrm{H_2}(g)}\left(T\right)/2 - b\mathrm{IHC}^{\circ}_{\mathrm{H_2O}(g)}\left(T\right) \tag{2.89}$$

We have included ZPE, entropy, and IHC changes in Eq. 2.84 because the adsorption of H_aO_b involves the consumption and production of molecules. E_{H_2} and E_{H_2O} are the DFT total energies of isolated H_2 and H_2O molecules, respectively. The ZPE of surfaces with and without adsorbates were calculated using DFPT where only the adsorbates and surface atoms near the binding site were considered. The ZPE, standard entropy S°, and IHC of H_2 and H_2O at STP were taken from thermochemical tables and reproduced in Table 2.1. (133; 143; 144) Note that the final three terms in Eq. 2.84 (*i.e.* those that depend on pH, U, and the vapor pressure) constitute a modified ΔG_{solv}, which is applicable when the dissolution or deposition of surface atoms does not accompany the adsorption or desorption of H_aO_b.

Using Eqs. 2.82 and 2.84, one can calculate the free energy changes associated with a plethora of aqueous surface chemical reactions under various conditions (of U and pH). These free energy changes can then be used to construct aqueous surface phase diagrams as follows:

1. Choose a reference surface.

2. Generate a catalog of surfaces that offer a wide variety of compositions and structures. This can be achieved by systematically removing atoms from and adding atoms to the reference surface (see Fig. 2.5).

3. Calculate the free energy differences between the surfaces and the reference as a function of U and pH.

4. Assign a color to each surface and, for each U and pH, plot the color of the surface with the smallest free energy difference.

Fig. 2.5 shows a schematic aqueous surface phase diagram of A_xB_y.

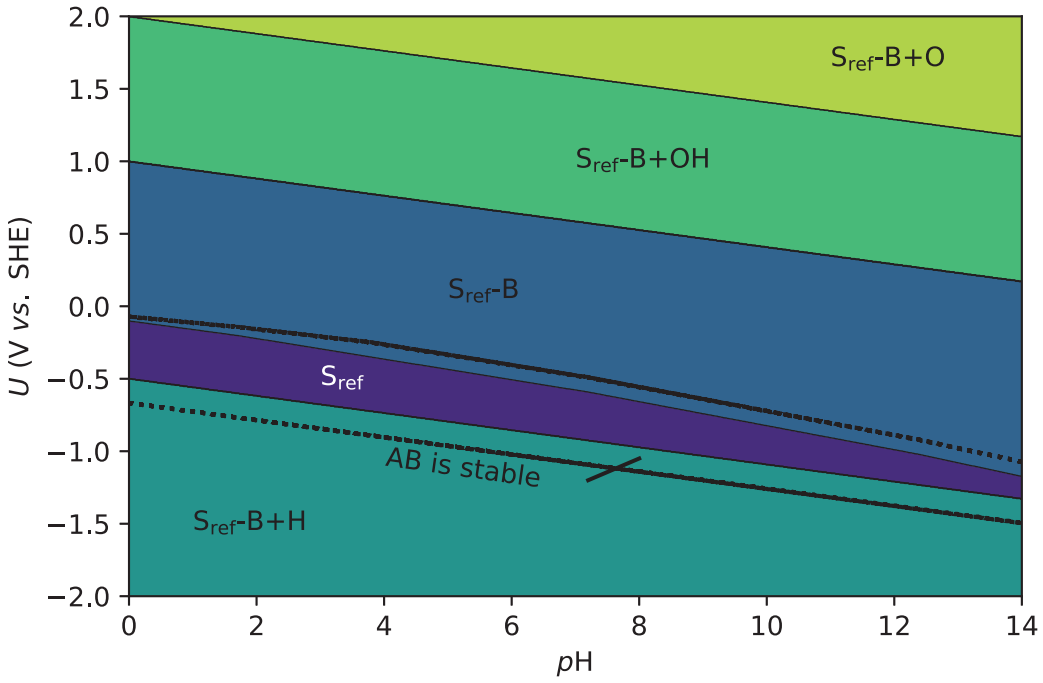

Figure 2.5: Aqueous surface phase diagram of a hypothetical binary compound A_xB_y. Five surface phases are shown, three of which (dark green, purple, and blue) are stable in the region of U and pH where AB is stable. The composition of the surface layer is defined relative to that of some reference (ref) S_{ref}, *e.g.* S_{ref}-B+H corresponds to a surface that has one B removed and one H adsorbed. Dotted lines correspond to the region of U and pH where AB is stable.

	ZPE	TS°	IHC
H_2	0.27	0.40	0.09
H_2O	0.56	0.58	0.10

Table 2.1: ZPE, standard entropy, and integrated heat capacity of H_2 and H_2O at 298 K and 1 bar in units of eV.

2.3. Grand canonical Monte Carlo

Up to this point, we have demonstrated how DFT calculations and *ab initio* thermodynamics allow the prediction of the preferred surface phases under different experimental conditions. Accurate prediction of surface phase diagrams is crucial for the accurate modeling of surface chemical reactions because adsorption and activation free energies depend strongly on the surface sites present. One of the greatest determinants of the accuracy of a surface phase diagram is the completeness of the catalog of surfaces used in its construction. As described above in section 2.2.4, the traditional way to generate this catalog is using a systematic search of atom additions and removals guided by chemical intuition. This set of surfaces, however, is biased by human-selection and therefore not guaranteed to include the preferred surfaces in nature. An alternative is to generate this catalog randomly. The problem with a random search is that the phase space of all possible surfaces is infinite and therefore unguided exploration is inefficient. In order to overcome the note deficiencies of these two approaches, we combine DFT calculations and grand canonical Monte Carlo (GCMC) simulations where the latter enables both the thermodynamic and stochastic guidance through surface phase space necessary to construct an accurate surface phase diagram.

2.3.1. Grand canonical ensemble

The concept of a grand canonical ensemble comes from the statistical mechanical treatment of a system in contact with a heat and particle reservoir. In other words, the system can exchange both heat and particles with its surroundings. Fig. 2.6 shows a schematic of the system plus reservoir. For the combined system, the total energy

$E^{(0)}$ and number of particles $N^{(0)}$ are fixed

$$E^{(0)} = E + E' \tag{2.90}$$

$$N^{(0)} = N + N' \tag{2.91}$$

where E (E') and N (N') are the energy and number of particles in the system (reservoir), respectively. If the system is in a state r, such that $E = E_r$ and $N = N_r$, then the number of states accessible to the combined system Ω is equal to the number of states accessible to the reservoir Ω', *i.e.*

$$\Omega\left(E^{(0)}, N^{(0)}\right) = \Omega'\left(E', N'\right) = \Omega'\left(E^{(0)} - E_r, N^{(0)} - N_r\right) \tag{2.92}$$

where the right-hand side was obtained using Eqs. 2.90 and 2.91. From this, it follows that the probability of the system being in state r is

$$P_r = c\Omega'\left(E^{(0)} - E_r, N^{(0)} - N_r\right) = c\exp\left[\ln\Omega'\left(E^{(0)} - E_r, N^{(0)} - N_r\right)\right] \tag{2.93}$$

where c is a normalization constant. Since the system is much smaller than the combined system, $E_r << E^{(0)}$ and $N_r << N^{(0)}$, we can expand $\ln\Omega'$ in a Taylor series about $E^{(0)}$ and $N^{(0)}$

$$\ln\Omega'\left(E^{(0)} - E_r, N^{(0)} - N_r\right) = \ln\Omega'\left(E^{(0)}, N^{(0)}\right) - \beta E_r - \alpha N_r \tag{2.94}$$

where

$$\beta = \frac{1}{k_{\mathrm{B}}T} = \left.\frac{\partial \ln\Omega'}{\partial E'}\right|_{E^{(0)}} \tag{2.95}$$

and

$$\alpha = -\beta\mu = \left.\frac{\partial \ln \Omega'}{\partial N'}\right|_{N^{(0)}} \tag{2.96}$$

Eqs. 2.95 and 2.96 stem from the fundamental thermodynamic relation, which states that

$$dU = TdS - pdV + \mu dN \tag{2.97}$$

where U is the internal energy. For constant V and N, it is simple to show that

$$\left(\frac{\partial S}{\partial U}\right)_{V,N} = k_{\mathrm{B}} \left(\frac{\partial \ln \Omega}{\partial U}\right)_{V,N} = \frac{1}{T} \tag{2.98}$$

where we have inserted the statistical definition of entropy, $S = k_{\mathrm{B}} \ln \Omega$. On the other hand, for constant U and V, we have

$$\left(\frac{\partial S}{\partial N}\right)_{U,V} = k_{\mathrm{B}} \left(\frac{\partial \ln \Omega}{\partial N}\right)_{U,V} = -\frac{\mu}{T} \tag{2.99}$$

Inserting Eqs. 2.94 and 2.96 into 2.93 results in

$$P_r = ce^{-\beta(E_r - \mu N_r)} \tag{2.100}$$

where c is obtained from the normalization condition

$$\sum_r P_r = c \sum_r e^{-\beta(E_r - \mu N_r)} = 1 \tag{2.101}$$

$$c = \frac{1}{Z} = \frac{1}{\sum_r e^{-\beta(E_r - \mu N_r)}} \tag{2.102}$$

and Z is called the partition function. The collection of states whose probabilities are given by the distribution function P_r is called the grand canonical ensemble.

If the state r corresponds to a particular set of particle coordinates ($\mathbf{r}^N$) and momenta ($\mathbf{p}^N$), then we can make the following substitutions:

$$E_r \to E\left(\mathbf{r}^N, \mathbf{p}^N\right) = \mathcal{V}\left(\mathbf{r}^N\right) + T\left(\mathbf{p}^N\right) = \mathcal{V}\left(\mathbf{r}^N\right) + \sum_{i=1}^{N} p_i^2/2m_i \tag{2.103}$$

$$N_r \to N \tag{2.104}$$

$$\sum_r \to \sum_{N=0}^{N^{(0)}} \frac{1}{N!h^{3N}} \int d\mathbf{r}^N \int d\mathbf{p}^N = \sum_{N=0}^{\infty} \frac{1}{N!h^{3N}} \int d\mathbf{r}^N \int d\mathbf{p}^N \tag{2.105}$$

Here, E is the Hamiltonian, $\mathcal{V}$ is the potential energy, and T is the kinetic energy. The sum over N includes all possible arrangements of $N^{(0)}$ particles in the system and reservoir where we take the limit as $N^{(0)} \to \infty$. Making these substitutions, Z becomes

$$Z\left(\mu, V, T\right) = \sum_{N=0}^{\infty} \frac{e^{\beta\mu N}}{N!h^{3N}} \int d\mathbf{r}^N e^{-\beta\mathcal{V}\left(\mathbf{r}^N\right)} \prod_{i=1}^{N} \int d\mathbf{p}_i e^{-\frac{\beta p_i^2}{2m_i}} \tag{2.106}$$

If the particles have the same mass m, then the product of the integrals over $\mathbf{p}$ has a compact solution, that is,

$$\prod_{i=1}^{N} \int d\mathbf{p}_i e^{-\frac{\beta p_i^2}{2m_i}} = \frac{h^{3N}}{\Lambda^{3N}} \tag{2.107}$$

where Λ is the thermal de Broglie wavelength,

$$\Lambda = \sqrt{h^2/2\pi m k_{\mathrm{B}} T} \tag{2.108}$$

As a final simplification, one can replace $\mathbf{r}^N$ with fractional coordinates ($\mathbf{s}^N$), which yields

$$Z\left(\mu, V, T\right) = \sum_{N=0}^{\infty} \int d\mathbf{s}^N \frac{e^{\beta\mu N} V^N}{N!\Lambda^{3N}} e^{-\beta\mathcal{V}\left(\mathbf{s}^N\right)} \tag{2.109}$$

where the volume V should not be confused with the potential energy $\mathcal{V}$. From Eq.

2.109, it is clear that the probability density corresponding to N particles at $\mathbf{s}^N$ is

$$\rho\left(\mathbf{s}^N; N\right) = \frac{e^{\beta\mu N} V^N}{N!\Lambda^{3N}} e^{-\beta\mathcal{V}\left(\mathbf{s}^N\right)} \tag{2.110}$$

This is the main statistical mechanical ingredient for GCMC.

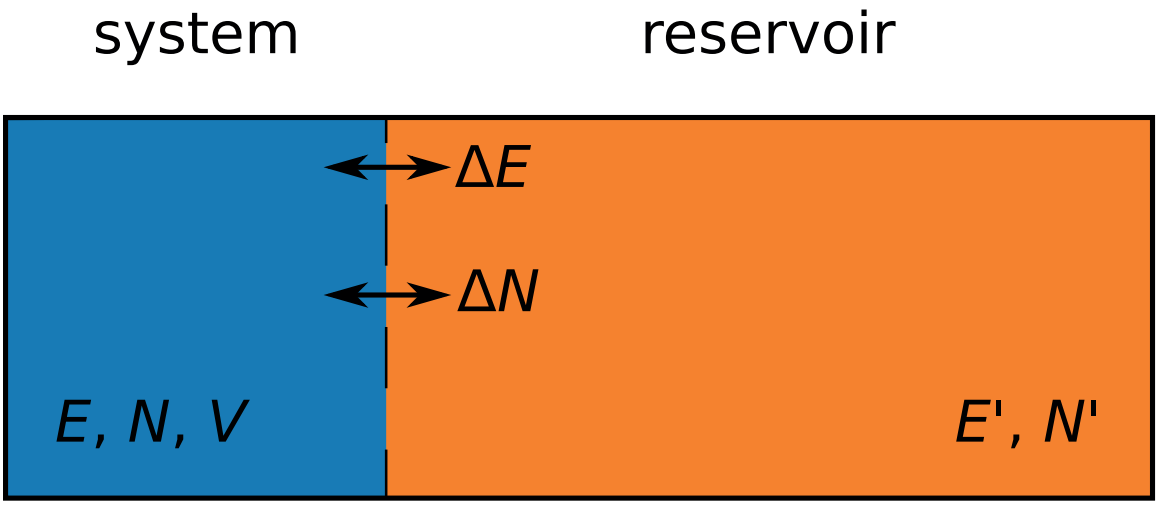

Figure 2.6: Schematic of the grand canonical ensemble. The system (with energy E, number of particles N, and volume V) is in contact with a reservoir (with energy E' and number of particles N') with which it can exchange heat ΔE and particles ΔN.

2.3.2. Metropolis Monte Carlo

Our aim is to determine the equilibrium structure and composition of the surface without having to preselect a catalog of surfaces. In principle, this can be achieved by sampling the entire grand canonical ensemble and identifying the surface with the highest probability density. Such an approach, however, is inefficient because many of the phase points ($\mathbf{s}^N$; N) sampled will have low probability densities. Therefore, we seek a scheme that enables the system to jump between neighboring phase points with high probability densities. The most widely used scheme in the physical sciences is that of Metropolis. (145)

To illustrate this scheme, imagine a system that is in the state o but is considering a jump to the state n where we denote the probability of this transition as $\pi(o \to n)$. At equilibrium, we assume that the rates of jumping from o to n and *vice versa* are equal:

$$\rho(o)\,\pi(o \to n) = \rho(n)\,\pi(n \to o) \tag{2.111}$$

where $\rho(i)$ is the probability density of being at the phase point i. Eq. 2.111 is called the detailed balance condition. Transitions in Monte Carlo (MC) simulations generally involve two steps: (1) propose a trial jump and (2) decide whether to jump. Therefore, we can rewrite $\pi(o \to n)$ as

$$\pi(o \to n) = \alpha(o \to n)\,\mathrm{acc}(o \to n) \tag{2.112}$$

where $\alpha(o \to n)$ is the probability of attempting the trial jump from o to n and acc is the probability of accepting it. In the Metropolis scheme, α is usually chosen to be symmetric, *i.e.* $\alpha(o \to n) = \alpha(n \to o)$. As a result, the detailed balance condition

becomes

$$\frac{\text{acc}\,(o \to n)}{\text{acc}\,(n \to o)} = \frac{\rho\,(n)}{\rho\,(o)} \tag{2.113}$$

where ρ is replaced by Eq. 2.110. Eq. 2.113 is small when state o is favored, large when state n is favored, and one when neither are favored. In the first case, the trial jump will likely be rejected whereas, in the second, it will likely be accepted. In the third case, forward and backward jumps occur with equal probability.

There are three different types of trial jumps in GCMC simulations: particle displacements ($\mathbf{s}^N \to \mathbf{s}^{N\prime}$), particle additions ($\mathbf{s}^N \to \mathbf{s}^{N+1}$), and particle removals ($\mathbf{s}^N \to \mathbf{s}^{N-1}$). Inserting these phase points into Eq. 2.113 gives

$$\frac{\text{acc}\left(\mathbf{s}^N \to \mathbf{s}^{N\prime}\right)}{\text{acc}\left(\mathbf{s}^{N\prime} \to \mathbf{s}^N\right)} = e^{-\beta\left[\mathcal{V}\left(\mathbf{s}^{N\prime}\right)-\mathcal{V}\left(\mathbf{s}^N\right)\right]} \qquad \textit{displacement} \tag{2.114}$$

$$\frac{\text{acc}\left(\mathbf{s}^N \to \mathbf{s}^{N+1}\right)}{\text{acc}\left(\mathbf{s}^{N+1} \to \mathbf{s}^N\right)} = \frac{V}{(N+1)\,\Lambda^3} e^{-\beta\left[\mathcal{V}\left(\mathbf{s}^{N+1}\right)-\mathcal{V}\left(\mathbf{s}^N\right)-\mu\right]} \qquad \textit{addition} \tag{2.115}$$

$$\frac{\text{acc}\left(\mathbf{s}^N \to \mathbf{s}^{N-1}\right)}{\text{acc}\left(\mathbf{s}^{N-1} \to \mathbf{s}^N\right)} = \frac{N\Lambda^3}{V} e^{-\beta\left[\mathcal{V}\left(\mathbf{s}^{N-1}\right)+\mu-\mathcal{V}\left(\mathbf{s}^N\right)\right]} \qquad \textit{removal} \tag{2.116}$$

where the probabilities of accepting these trial jumps are

$$\text{acc}\left(\mathbf{s}^N \to \mathbf{s}^{N\prime}\right) = \min\left\{1, e^{-\beta\left[\mathcal{V}\left(\mathbf{s}^{N\prime}\right)-\mathcal{V}\left(\mathbf{s}^N\right)\right]}\right\} \qquad \textit{displacement} \tag{2.117}$$

$$\text{acc}\left(\mathbf{s}^N \to \mathbf{s}^{N+1}\right) = \min\left\{1, \frac{V}{(N+1)\,\Lambda^3} e^{-\beta\left[\mathcal{V}\left(\mathbf{s}^{N+1}\right)-\mathcal{V}\left(\mathbf{s}^N\right)-\mu\right]}\right\} \qquad \textit{addition} \tag{2.118}$$

$$\text{acc}\left(\mathbf{s}^N \to \mathbf{s}^{N-1}\right) = \min\left\{1, \frac{N\Lambda^3}{V} e^{-\beta\left[\mathcal{V}\left(\mathbf{s}^{N-1}\right)+\mu-\mathcal{V}\left(\mathbf{s}^N\right)\right]}\right\} \qquad \textit{removal} \tag{2.119}$$

where the min function returns the smaller value. For example, if the second argument of the min function is 0.5, then the trial jump is accepted with a probability of 0.5. Eqs. 2.117, 2.118, and 2.119 comprise the acceptance rules for GCMC simulations.

Fig. 2.7 shows a flow chart of a GCMC program. First, one defines the V of the system, the μ and T of the reservoir, an initial slab model, and configurational biases (if any). Then, the algorithm decides randomly whether to add, remove, or displace a particle. If it decides to add a particle, then it chooses a site. For combined systems containing two or more elements, choosing an element precedes choosing a site. The latter can be done randomly, however, the new particle may end up too close or too far from the other particles in the system. As such, configurational biases are often employed to prepare trial jumps to more realistic states, *e.g.* imposing minimum and maximum distances between the new and old particles. After preparing a trial jump, the total energy of the system is calculate, which, in the case of our *ab initio* GCMC, is done using DFT. Additionally, since DFT calculations of slabs can be computationally expensive, we remove displacements as a trial jump and replace the total energy calculation with a structural optimization. This enables us to focus on surface reconstruction processes, which generally involve atom additions and removals, while still accounting for the local relaxation that would normally occur during displacement steps. Once the total energies of the old and new structures have been calculated, they are inserted into Eqs. 2.117, 2.118, and 2.119, which determine whether the trial jump should be accepted. If it is rejected, then the previous structure is restored and the loop continues.

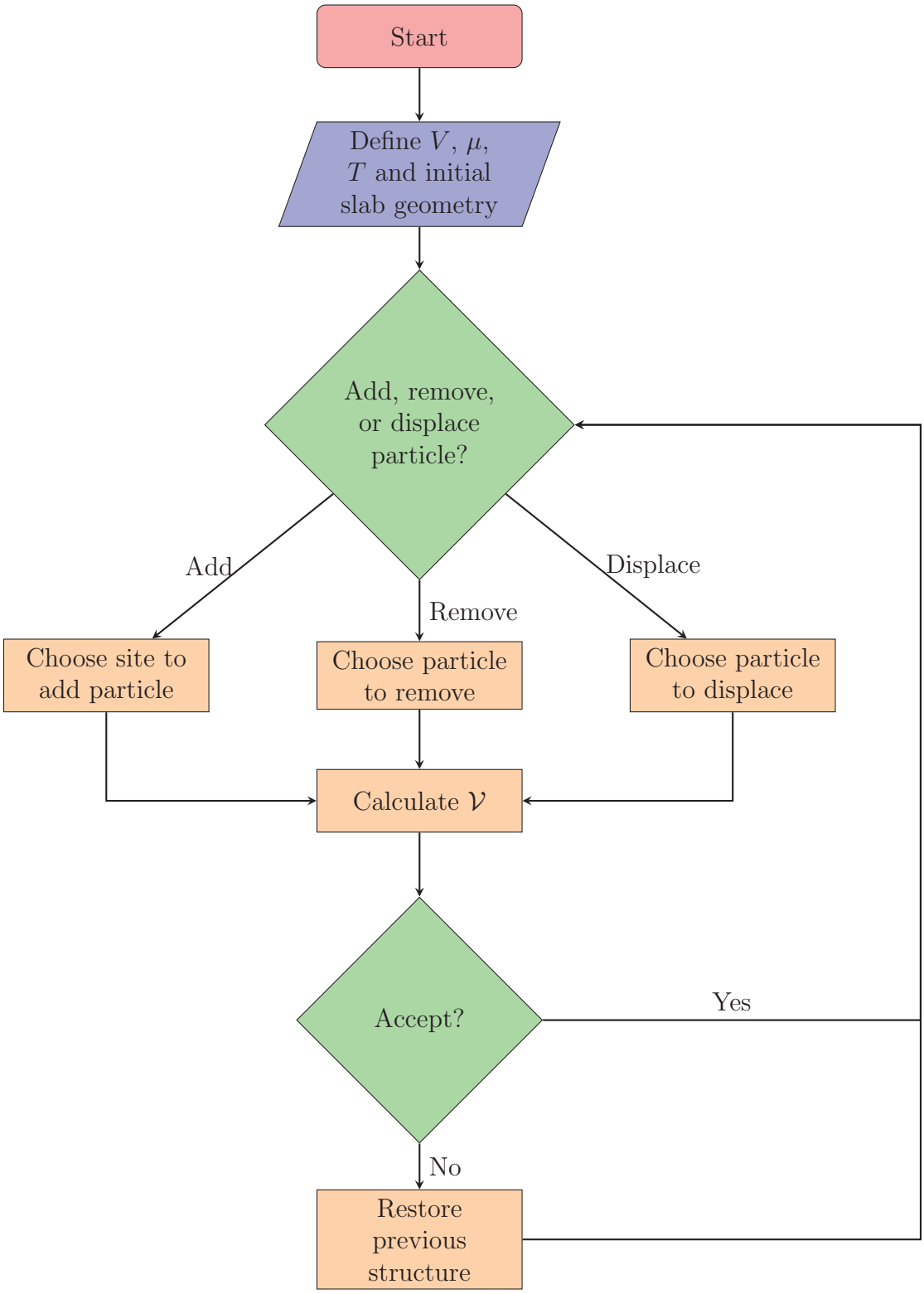

Figure 2.7: Flow chart of a GCMC simulation.

2.4. Tree-based methods for machine learning

Up to this point, we have explained how to accurately predict surface chemical reactivity using DFT, *ab initio* thermodynamics, and GCMC. We have not yet, however, explained how one might identify the structural and electronic origin of this reactivity and subsequently control it. Generally speaking, we wish to map the structural and electronic degrees of freedom (DOF) to properties of interest. The DOF in surface chemical reactions are the nuclear coordinates of the surface and adsorbates and the electronic wave function. These DOF provide a complete but unintuitive description of the surface. Fortunately, it is often possible to find a small set of intuitive DOF that, despite their simplicity, can still be mapped to surface chemical reactivity. For example, the nuclear coordinates can be replaced by local structural DOF like surface bond lengths, angles, and coordination numbers. Similarly, the electronic wave function can be substituted with local electronic DOF such as Bader charges. (146–149) These four DOF are but a small subset of those that are possible (for more examples, see Refs. 150, 151, and 152). Having identified a few intuitive DOF, we can now map them to quantities that determine surface chemical reaction rates, *e.g.* adsorption and activation free energies, using statistical learning methods. Of the many methods available, (153; 154) we use random forests (RFs) because they can capture complex nonlinearities in the data and yet they are easy to train, tune, and interpret.

2.4.1. Decision trees

Before growing RFs, one must first learn how to grow decision trees (DTs). DTs ask the series of yes/no questions based on the inputs x that best separate the data based on the response y. To see how this works, consider the data shown in Fig. 2.8.

Here, there are two inputs (x_1 and x_2), one response (see color bar for its magnitude), and 1,000 observations. In order to separate the data based on y, we might start by asking the question "is $x_1 \leq s_1$?". This question can be thought of as a vertical line at $x_1 = s_1$ that splits the data into two groups, one where $\bar{y} \approx 1.5$ (group A) and another where $\bar{y} \approx 4$ (group B). While this question is effective at coarsely separating the data based on y, we can achieve higher resolution by asking follow-up questions. For example, by asking "is $x_2 \leq s_3$?" on group A, we find two subgroups (purple and blue) with nearly constant y ($\bar{y} \approx 1$ and $\bar{y} \approx 2$, respectively). The process of asking follow-up questions on groups and subgroups continues until some stopping criteria are met. Typical stopping criteria include the maximum depth of the DT and the minimum number of observations in a group. A DT for this data is shown in Fig. 2.9.

Using this DT, let us now explain how to calculate the predicted response $\hat{y}$ for a new observation located at the center of Fig. 2.8, *i.e.* at $x_1 = (s_1 + s_2)/2$ and $x_2 = (s_3 + s_4)/2$. First, the DT asks the observation if its $x_1 \leq s_1$. Since the answer is "no", the observation moves down and to the right whereupon the DT asks if its $x_1 \leq s_2$. Since the answer is "yes", the DT returns $\hat{y} \approx 3$, which agrees with the color distribution in Fig. 2.8.

The algorithm for growing DTs can be generalized as follows. Consider a data set with p inputs, one response, and N observations. The observation i of this data set can be written as (x_i, y_i) where $x_i = (x_{i1}, \ldots, x_{ij}, \ldots, x_{ip})$. First, we ask a question on the data. More precisely, we specify a splitting variable j and split point s and then ask "is $x_j \leq s$?" This question splits the data into two groups, G_1 and G_2. If we model the response as a constant c_k in each group G_k, then the predicted response

for observation i is

$$\begin{aligned}\hat{y}_i &= \sum_{k=1}^{2} c_k I\left(x_i \in G_k\right) \\ &= c_1 I\left(x_i \in G_1\right) + c_2 I\left(x_i \in G_2\right)\end{aligned} \tag{2.120}$$

where I is a function that returns 1 if $x_i \in G_k$ and 0 otherwise. Now, we need to determine the best c_k, which can be done using the method of least squares. In this method, the best c_k are those that minimize the sum of squared residuals (SSR), *i.e.* those that solve

$$\frac{\partial}{\partial c_k} \sum_{i=1}^{N} \left(y_i - \hat{y}_i\right)^2 = \frac{\partial}{\partial c_k} \sum_{i=1}^{N} \left[y_i - \sum_{k=1}^{2} c_k I\left(x_i \in G_k\right)\right]^2 = 0 \tag{2.121}$$

where $y_i - \hat{y}_i$ is the residual for observation i. It is straightforward to show that the best c_k are

$$\hat{c}_k = \frac{\sum_{i=1}^{N} y_i I\left(x_i \in G_k\right)}{\sum_{i=1}^{N} I\left(x_i \in G_k\right)} = \text{ave}\left(y_i | x_i \in G_k\right) \tag{2.122}$$

where ave stands for average. Inserting Eq. 2.122 into 2.120 gives the least squares predicted response $\hat{y}_i^{\text{LS}}$ for observation i,

$$\begin{aligned}\hat{y}_i^{\text{LS}} &= \sum_{k=1}^{2} \hat{c}_k I\left(x_i \in G_k\right) = \sum_{k=1}^{2} \text{ave}\left(y_i | x_i \in G_k\right) I\left(x_i \in G_k\right) \\ &= \text{ave}\left(y_i | x_i \in G_1\right) I\left(x_i \in G_1\right) + \text{ave}\left(y_i | x_i \in G_2\right) I\left(x_i \in G_2\right)\end{aligned} \tag{2.123}$$

There are pN possible questions that we can ask, each corresponding to a unique (j, s) pair. We seek the question that minimizes the SSR. Once this question has been identified, we split the data and ask follow-up questions on the resulting groups until the DT is sufficiently deep.

If the DT is too deep, however, it risks overfitting the data. In other words, it

learns spurious relationships between the inputs and residuals. To combat this, we use cost-complexity pruning, which decreases model complexity by removing questions from the DT. For instance, imagine $\bar{y} \approx 4.1$ for $x_2 > s_4$ in Fig. 2.9. If the stopping criteria are not met, then the DT will ask if $x_2 \leq s_4$. Since the average responses of the resulting groups are similar (4 *vs.* 4.1), this question does not improve model performance and therefore can be removed.

Suppose that we grow a DT T_0 and then prune it to T where $T \subset T_0$. The cost of this action is

$$C_\alpha(T) = \sum_{k=1}^{|T|} N_k \text{MSE}_k(T) + \alpha|T| \tag{2.124}$$

where $|T|$ is the number of groups in T, N_k is the number of observations in G_k, $\text{MSE}_k(T)$ is the mean squared error of y in G_k of T, and α is a parameter that controls model complexity. Our goal is to find the subtree T_α that minimizes C_α for each α. In general, T_α is small when α is large, T_α is large when α is small, and $T_\alpha = T_0$ when $\alpha = 0$. The best T_α can be found using weakest link pruning, which consecutively removes the question that incurs the smallest per-question increase in the SSR until only one question remains. It has been shown that the sequence of DTs generated by weakest link pruning must contain T_α. (155; 156) The final step is to determine the α that minimizes the cross-validated SSR.

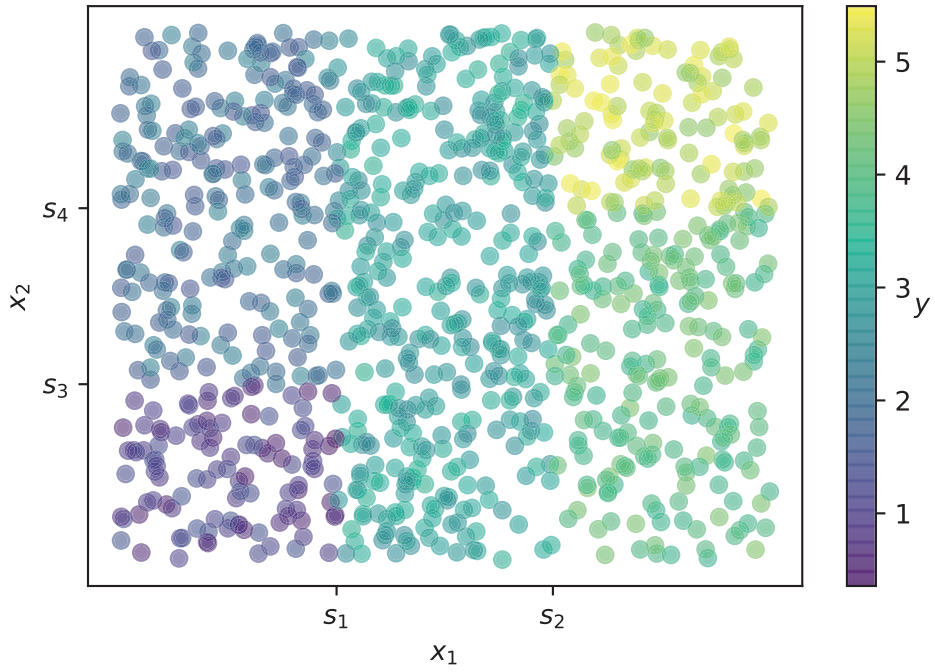

Figure 2.8: Data used to grow the DT in Fig. 2.9. The data consists of two inputs x_1 and x_2, one response y, and 1000 observations. The color of the observations corresponds to y.

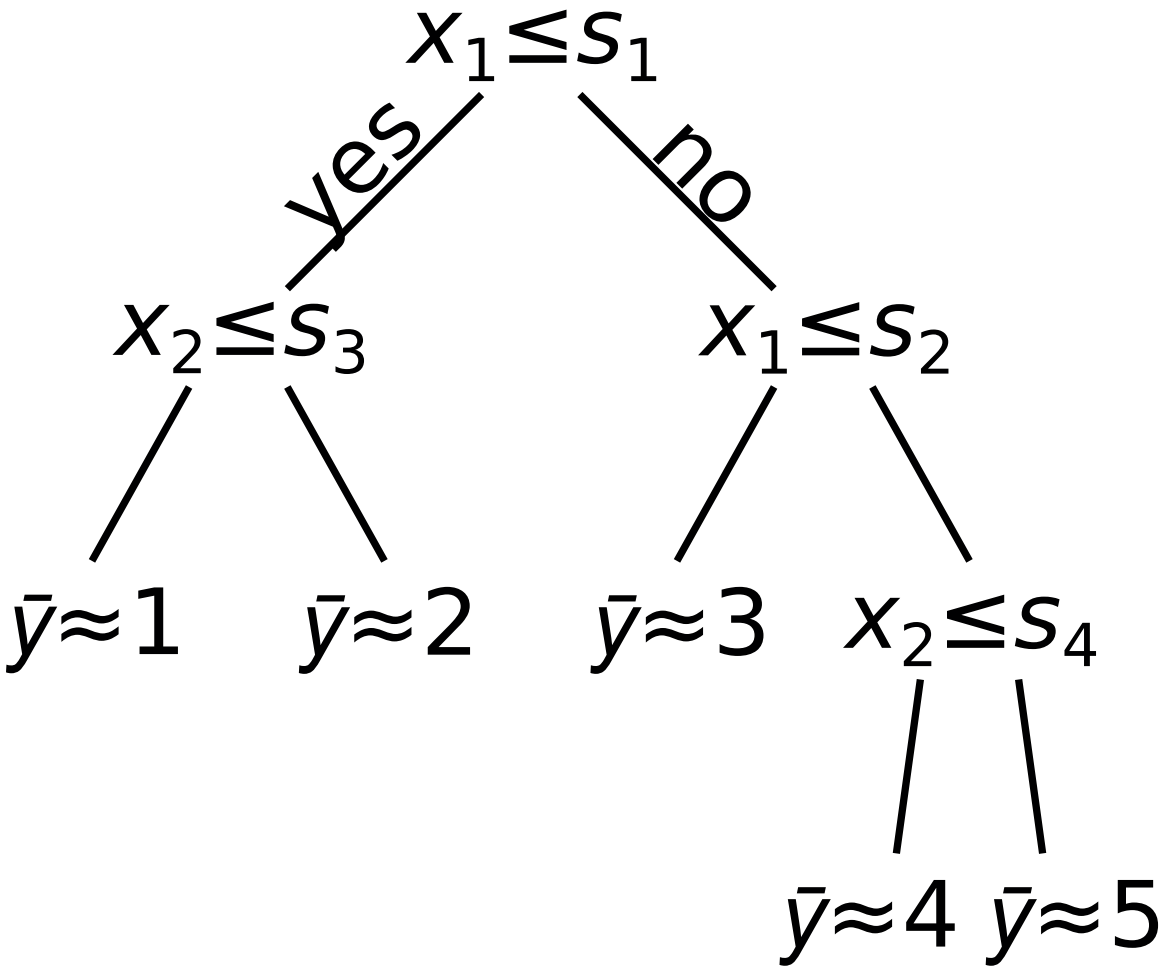

Figure 2.9: DT for the data shown in Fig. 2.8. Questions (inequalities) split the data into groups, which are further split into subgroups by follow-up questions. Predictions are made by passing new observations from question to question until they are assigned a $\bar{y}$.

2.4.2. Random forests

One major drawback of DTs is their tendency to overfit data due to model complexity. Overfitting can be remedied to some extent by growing shallower trees and pruning them. This, however, comes at the expense of model performance. A relatively simple way to rectify overfitting whilst preserving model performance is to generate an ensemble of trees trained on different parts of the data. Such models are called RFs, an example of which is shown in Fig. 2.10. The procedure of growing this RF can be summarized as follows. First, we prepare a data set (x_{ij}, y_i) where i and j index the observations and inputs, respectively. Then, we randomly select a subset of the observations and inputs, and use them to grow a DT. This step is repeated n times and has the effect of lowering model variance (*i.e.* improving its ability to make accurate predictions on new observations) without significantly reducing model complexity. Since each DT is trained on a different parts of the data, they acquire different shapes, *e.g.* different depths, numbers of question (represented as circles), and branching patterns. This results in a RF containing a diverse set of DTs. Using this RF, we can then make predictions on a new observation by passing it through each of the n DTs (see blue circles connected by arrows). Each DT predicts a slightly different value for the response $\{\hat{y}_{i1}, \ldots, \hat{y}_{ik}, \ldots, \hat{y}_{in}\}$, which are then averaged to give the final prediction $\hat{y}$. This averaging serves to decrease model variance by weakening the impact of an overfitted DT on $\hat{y}$.

Another way to rectify overfitting is regularization, which rewards the growth of models with fewer inputs. It therefore has the added benefit of increasing model interpretability. In the RF algorithm, regularization can be used to modify the selection of inputs from which DTs are grown. The details of its implementation will now be discussed. (157–159) First, however, we need to introduce the concept of gain.

Recall that, when growing DTs, we ask many questions and choose the one that minimizes the SSR. Gain is simply defined as the magnitude of the reduction in the SSR upon asking a question. Therefore, another way to think of minimizing the SSR is maximizing the gain. Now, we return to the growth of the regularized RFs (RRFs). The first step is to grow a DT in the usual way. We denote the set of inputs used as splitting variables by F. Then, we proceed to grow a second DT. This time, however, we penalize the use of inputs not in F by reducing their gain as follows:

$$\text{gain}_R\left(x_j\right) = \begin{cases} \lambda \cdot \text{gain}\left(x_j\right) & x_j \notin F \\ \text{gain}\left(x_j\right) & x_j \in F \end{cases} \tag{2.125}$$

Here, the subscript R stands for regularized and $\lambda \in [0, 1]$ is the regularization coefficient. If $\lambda = 1$, then a RF is grown from all the inputs. On the other hand, if λ is smaller, then a RRF is grown from a subset of the inputs, *i.e.* those in F. For the remaining $n - 2$ DTs in the RRF, F is updated to include new inputs x_{new} for which $\text{gain}_R\left(x_{\text{new}}\right) > \max\left[\text{gain}_R\left(\{x|x \in F\}\right)\right]$. For more details on DTs and RFs, the reader is referred to the Hastie's classic text on statistical learning. (160)

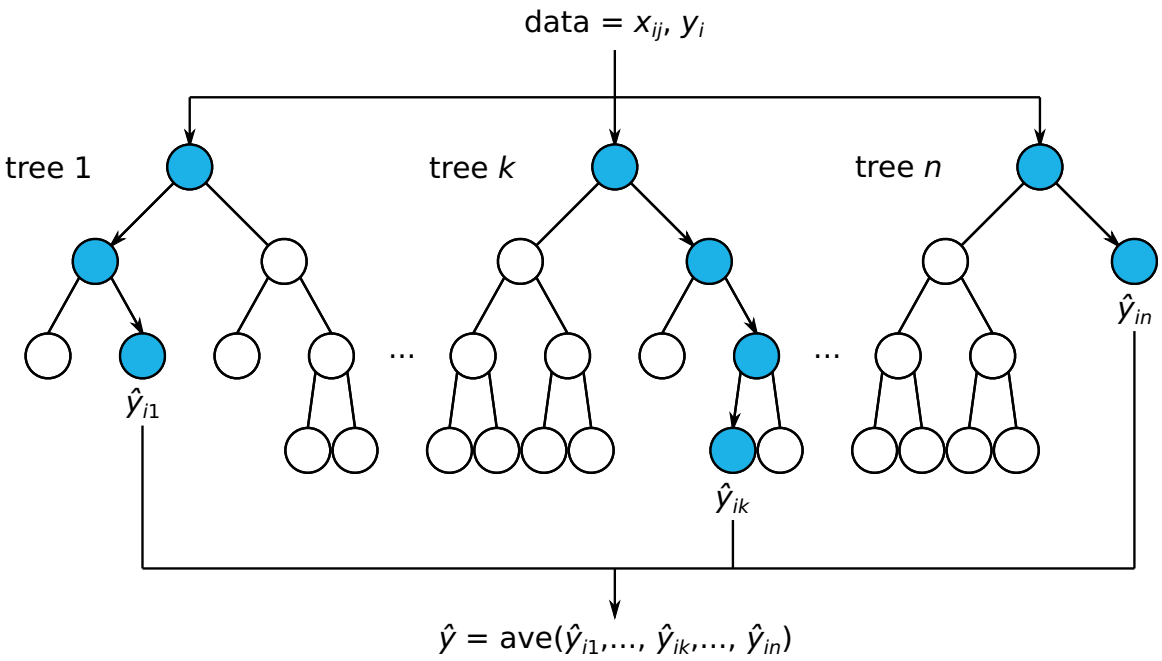

Figure 2.10: Schematic of a RF, which consists of n DTs trained on different subsets of the observations i and inputs j. Circles correspond to questions. The predicted response $\hat{y}$ of a new observation is calculated by passing it through each DT, shown as blue circles connected by arrows, and averaging their responses.

CHAPTER 3 : Phosphorus-decorated reconstructions of the (0001) surface of Ni_2P and Ni_5P_4

3.1. Introduction

Elemental enrichment of the surfaces of transition metal oxides, sulfides, and phosphides is a desirable route to modify their physical and/or chemical properties. The ability to manipulate the amount of the transition metal component of a transition metal compound is believed to be the key to modifying the chemistry of its surface. Perhaps for oxides this is naively rooted in the fact that oxygen cannot stably form extended compounds or chains exceeding three atoms (*e.g.* ozone), and thus surfaces of oxides are most directly manipulated through their metallic components. Phosphorus and sulfur, however, can form extended chains or crystals on their own; thus P and S present possible sources of solid secondary phases when they form compounds with metals. (161–163) This behavior may have deeper implications for the types of stable surfaces metal phosphides or sulfides form.

Surface composition and structure govern the key catalytic properties of nickel phosphides for a number of industrially important chemical reactions. (32; 34; 164–166) The material properties of transition metal phosphides, including binary, ternary, and mixed-metal compounds, have been well characterized since the 1960s. (167; 168) They form binary phosphides M_xP_y (where M denotes a transition metal) with compositions usually ranging from M_3P to MP_3 including Ni_2P, Ni_5P_4, Ni_3P, Ni_5P_2, $Ni_{12}P_5$, NiP, and NiP_2.

The metal-rich phosphides (M_xP_y where $x > y$) tend to exhibit better thermal stability, chemical inertness, electrical conductivity (169–171), and hardness than P-

rich phases; therefore we focus on Ni_2P and Ni_5P_4 in this study. Both solids have hexagonal crystal structures, as shown in Fig. 3.1. Their low index (0001) facets are found to be the most stable terminations for both compounds. Ni_2P has two unique terminations on the (0001) plane, the Ni_3P and Ni_3P_2 layers, and they have distinct structure and stoichiometry. DFT calculations reveal that the latter is more stable under conditions where bulk Ni_2P is stable. (172) Calculated scanning tunneling microscope (STM) images for this termination (172) do not, however, match experimental data. (171; 173; 174) This apparent inconsistency was resolved by dynamic low-energy electron diffraction (LEED) experiments, which revealed a surface reconstruction where non-stoichiometric additional P covers the Ni_3P_2 surface at the Ni_3-hollow sites (see Fig. 5 in Ref. 175). They showed that this P-covered Ni_3P_2 termination (hereafter denoted as Ni_2P-Ni_3P_2+P) comprises $\approx$ 80% of the surface, whereas the Ni_3P_2 (hereafter denoted as Ni_2P-Ni_3P_2) makes up the remaining 20%. Other reconstructions have been reported but only after annealing at higher temperatures, *e.g.* above 790 K. (173; 176) By contrast, Ni_5P_4 has a larger unit cell composed of Ni_3P_2, Ni_3P_3 and Ni_4P_3 layers (Fig. 3.1b). Additionally, Ni_5P_4 lacks reflection symmetry with respect to the [0001] plane. The surface structures of Ni_5P_4 have not been studied previously. Although both are demonstrated to be efficient catalysts for H_2 evolution, Ni_5P_4 is much more active than Ni_2P, thereby igniting interest in the differences and similarities of these surfaces with the hope of elucidating their hydrogen evolution reaction (HER) catalytic activity. (34)

In this chapter we conduct a comprehensive, systematic theoretical investigation of surface reconstruction on the (0001) and (000$\bar{1}$) facets of Ni_2P and Ni_5P_4. First, we calculate the standard energy of formation for different bulk nickel phosphides (Ni_xP_y) with first principles calculations so as to determine the region of P chemical potential

($\Delta\mu_P$) where bulk Ni_2P and Ni_5P_4 are stable. Then, we compute the free energy for a large number of distinct surface terminations. We show that both $Ni_2P(0001)$ and $Ni_5P_4(0001)$ prefer chemical and structural reconstructions of their bulk terminations, thereby providing a framework for the re-examination of the HER mechanism on these nickel phosphides.

3.2. Computational Methods

DFT (72; 177) calculations with periodic boundary conditions are performed using the Quantum ESPRESSO code. (178) The generalized gradient approximation (GGA) is employed to calculate the exchange-correlation energy using the method of Perdew, Burke, and Ernzerhof (PBE). (87) Optimized (106), norm-conserving, designed non-local (105) pseudopotentials were generated using OPIUM (179) to represent the core electrons and soften the potential for the valence electron wavefunctions. Semicore $3p$ states are treated along with the $4s$, $4p$, and $3d$ for Ni, while $3s$, $3p$, and $3d$ valence states are included for P. The $3d$ orbitals of P are deemed necessary to model anionic electron configurations and complex coordination environments. The valence states are expanded in a plane-wave basis with a 50 Ry kinetic energy cutoff. For bulk calculations, 5×5×6 and 5×5×4 k-point grids, shifted along k_z, are used to sample the Brillouin zones of Ni_2P and Ni_5P_4 respectively, and a small (0.005 Ry) Gaussian electronic smearing is applied to improve electronic convergence. All cells are allowed to fully relax, with stress convergence threshold of 0.01 kbar. We perform spin-polarized calculations for bulk Ni_2P and Ni_5P_4. The spin configuration is initially ferromagnetic, but both systems relax to a non-ferromagnetic state. Orbital-projected density of states (DOS) calculations and Löwdin population analysis are performed on the stable surface reconstructions of Ni_2P and Ni_5P_4 using a finer k-point grid (12×12×1).

Surfaces are modeled using slabs with 1×1 (for high surface defect concentrations) and $\sqrt{3}\times\sqrt{3}\ R30^\circ$ (for fractional defect concentrations) surface cells separated by $\approx$ 15 Å of vacuum. For slabs, the k-point grid sampling is reduced to 5×5×1 and 3×3×1 respectively. A dipole correction (180) is applied to further remove the artificial electric field interactions between periodic images; its contribution to the total energy, however, is very minor. We consider the (0001) surface of both Ni_2P and Ni_5P_4. Seven-layer, up-down symmetric slabs are used to model Ni_2P(0001) terminations. For Ni_5P_4, asymmetric slabs of thickness ranging from six to eight layers are constructed. In studying the reconstructions of the (0001) surface of Ni_5P_4, the composition of the $(000\bar{1})$ is kept to be Ni_4P_3. Likewise, in studying $(000\bar{1})$, the same type of termination is kept for (0001). All atoms are allowed to fully relax with force convergence threshold of 10^{-3} Ry/Bohr.

The surface free energy (Ω) is calculated under various environmental conditions using the following expression:

$$\Omega = \frac{1}{2A}(\phi + \Gamma_{\mathrm{P}}\Delta\mu_{\mathrm{P}}) \tag{3.1}$$

where A is the area of the surface unit cell, $\Delta\mu_{\mathrm{P}}$ is the chemical potential of phosphorus relative to bulk phase white phosphorus, and ϕ and Γ_{P} are

$$\phi = E_{\mathrm{slab}} - \frac{N_{\mathrm{Ni}}E^{\mathrm{bulk}}_{\mathrm{Ni}_x\mathrm{P}_y}}{x} + \Gamma_{\mathrm{P}}E^{\mathrm{bulk}}_{\mathrm{P}} \tag{3.2}$$

$$\Gamma_{\mathrm{P}} = N_{\mathrm{Ni}}(y/x) - N_{\mathrm{P}} \tag{3.3}$$

Here, E_{slab} is the DFT total energy of the slab, $N_{\mathrm{Ni/P}}$ is the number of Ni/P atoms in the slab, x/y is the number of Ni/P atoms per formula unit of the bulk, and E^{bulk} is the formation energy of the bulk compound in the subscript. Eq. 3.1 is governed

by the equilibrium conditions defined by the dominant bulk phase and a secondary P chemical reservoir.

3.3. Results and Discussion

3.3.1. Bulk Stability of Ni_2P and Ni_5P_4

A prerequisite of the existence of a surface is the presence of a stable bulk phase that supports it. Thus, to be able to determine the chemical potential ranges where certain surfaces become relevant, we generate a bulk phase diagram from first principles thermodynamic calculations (Fig. 3.1c). Another P-rich phase, NiP, is expected to dominate in the same P chemical potential range as NiP_2. However, NiP was demonstrated to only to appear at high temperatures (> 850 K) (181) and thus not included in the phase diagram (Fig. 3.1c) which represent only low temperature conditions since we neglect entropy contributions. It shows the values of $\Delta\mu_{\mathrm{P}}$ and $\Delta\mu_{\mathrm{Ni}}$ where bulk Ni_2P and Ni_5P_4 are stable compared to other nickel phosphide compositions. We find that Ni_2P is stable at -0.78 eV $\leq \Delta\mu_{\mathrm{P}} \leq$ -0.50 eV. At $\Delta\mu_{\mathrm{P}}$ = -0.50 eV, bulk Ni_5P_4 becomes thermodynamically favored up to $\Delta\mu_{\mathrm{P}}$ = -0.28 eV, at which point bulk NiP_2 becomes favored. The regions where these two phases are stable define the chemical potentials where their respective surfaces and their corresponding reconstructions are bound to exist. These surface reconstructions are discussed in the following sections.

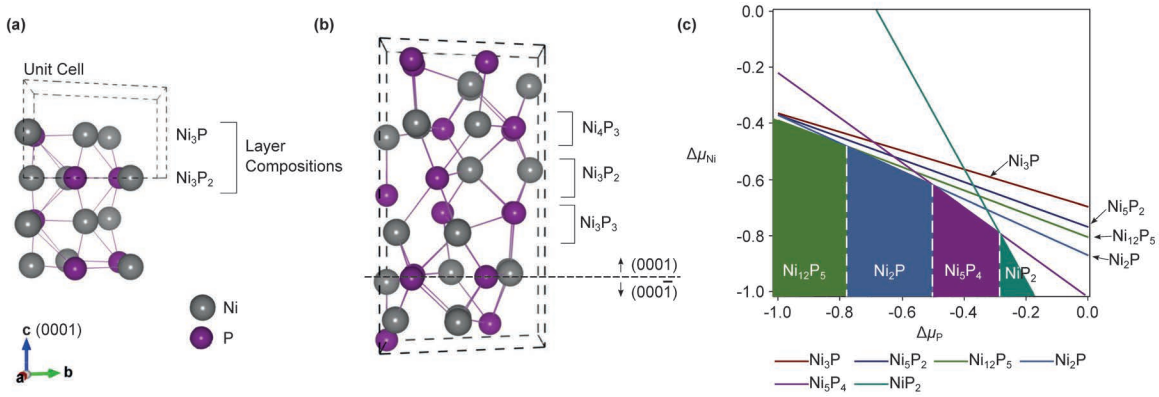

Figure 3.1: Crystal structure of (a) Ni_2P and (b) Ni_5P_4, highlighting the layer stacking along the [0001] direction. (c) Bulk phase diagram as a function of the relative chemical potentials (eV) of Ni and P for nickel phosphide phases that are stable under 850 K as has been found in the literature. Each solid line satisfies the equilibrium equation: $\Delta G_{\mathrm{Ni}_x\mathrm{P}_y} = x\Delta\mu_{\mathrm{Ni}} + y\Delta\mu_{\mathrm{P}}$ at 0 K. Each dotted vertical line denotes a phase coexistence point.

3.3.2. Surface Structure and Stability of $Ni_2P(0001)$

First, we describe the bulk-derived terminations of $Ni_2P(0001)$. As shown in Fig. 3.1a, bulk Ni_2P has two distinct layer compositions along the *c*-axis, Ni_3P and Ni_3P_2, where two layers constitute a unit cell. The surface crystal structure of Ni_2P-Ni_3P is shown in Fig. 3.2a for a $\sqrt{3} \times \sqrt{3}$ $R30°$ surface unit cell. The outermost layer consists of repeating Ni_3P subunits (green shaded triangle) with the central P sticking slightly out of the surface plane (shown in the inset). This surface also possesses six unique Ni_3 hollow sites (red shaded circles), which are formed by the Ni corners of the Ni_3P subunits. Ni_3P subunits are also found on Ni_2P-Ni_3P_2 but adopt a geometry where the central P is nearly coplanar with the Ni corners (see Fig. 3.2b). Owing to the increased surface P concentration on this termination, the number of unique Ni_3 hollow sites reduces to three. The center of the Ni_3P subunits in the Ni_2P-Ni_3P layer are above (and below) the Ni_3 hollows of the Ni_2P-Ni_3P_2, and *vice versa*. This stacking arrangement reveals a Ni_3P-Ni_3 pattern along the [0001] direction. The surface energy of Ni_2P-Ni_3P and Ni_2P-Ni_3P_2 is plotted in Fig. 3.2c as a function of $\Delta\mu_P$. In bulk Ni_2P stability region, Ni_2P-Ni_3P_2 is the preferred bulk-like termination, which is consistent with previous DFT calculations in the literature. (172)

Given that there is experimental evidence for P-enrichment of Ni_2P surfaces (171; 173–175; 182), we investigate the possibility of stabilizing bulk terminations by systematically adding extra P to the surface and then identify the most stable reconstruction. We find that the adsorption of P at Ni_3 hollow sites in a trigonal pyramidal geometry decreases the surface free energy of both Ni_2P-Ni_3P and Ni_2P-Ni_3P_2 (see Fig. 3.2c). This particular binding site and geometry is expected because it is analogous to creating a new, partial layer with P at bulk lattice positions. It is also different than the Ni_3P subunit, which has a nearly planar geometry. The Ni-rich (P-

poor) Ni_2P-Ni_3P termination is maximally stabilized by the addition of P at five out of the six Ni_3 hollow sites to generate a surface composition of Ni_3P+(5/3)P (defect concentration denoted per 1×1 surface unit cell). Higher coverages through further adsorption of P at the sixth Ni_3 hollow site (Ni_2P-Ni_3P+2P) causes the surface free energy to slightly increase, signaling P saturation. For Ni_2P-Ni_3P_2, the surface energy is minimized by the addition of P at all three Ni_3 hollow sites (Ni_3P_2+P in Fig. 3.2d). If a fourth P is added to make Ni_2P-Ni_3P_2+(4/3)P, it binds to a Ni-P bridge site, forms a surface adsorbed P_2 complex with P from the nearest Ni_3 hollow site (see Fig. A.1 in Appendix A), and increases the surface energy. Comparing the surface energies of both bulk-like terminations and their reconstructions, Ni_2P-Ni_3P_2+P is the most stable surface phase whenever bulk Ni_2P is stable, *i.e.* under relatively P-rich conditions. Thus, for a thermodynamically controlled synthesis of Ni_2P, this surface should be the most prevalent phase and, in fact, it is the same structure that Ref. 175 predicts, which they found to be the most consistent with the experimental STM among the surfaces they investigated.

The presence of P adatoms in Ni_3-hollow sites on Ni_2P-Ni_3P_2+P has important implications for the catalytic activity of $Ni_2P(0001)$ toward HER. It was proposed that H can adsorb at both Ni_3-hollow and Ni-P bridge sites, where the latter offers more moderate binding. (33) H atoms at these two sites can diffuse to one another and subsequently desorb, forming $H_2(g)$. Adlayer P offers a new site for H adsorption while also blocking the conventional bulk-like Ni_3 sites, thereby prompting a reconsideration of the HER mechanism on $Ni_2P(0001)$. (183) Ni_3-hollow sites play a key role in the predicted mechanisms of other important chemical reactions such as hydrodesulfurization (HDS) (184; 185), water-gas shift (WGS) (165), and hydrodeoxygenation (HDO). (186) Their catalytic mechanisms should be revisited

on reconstructed $Ni_2P(0001)$ surfaces since the above analysis shows that Ni_3-hollow sites may not in fact be readily available.

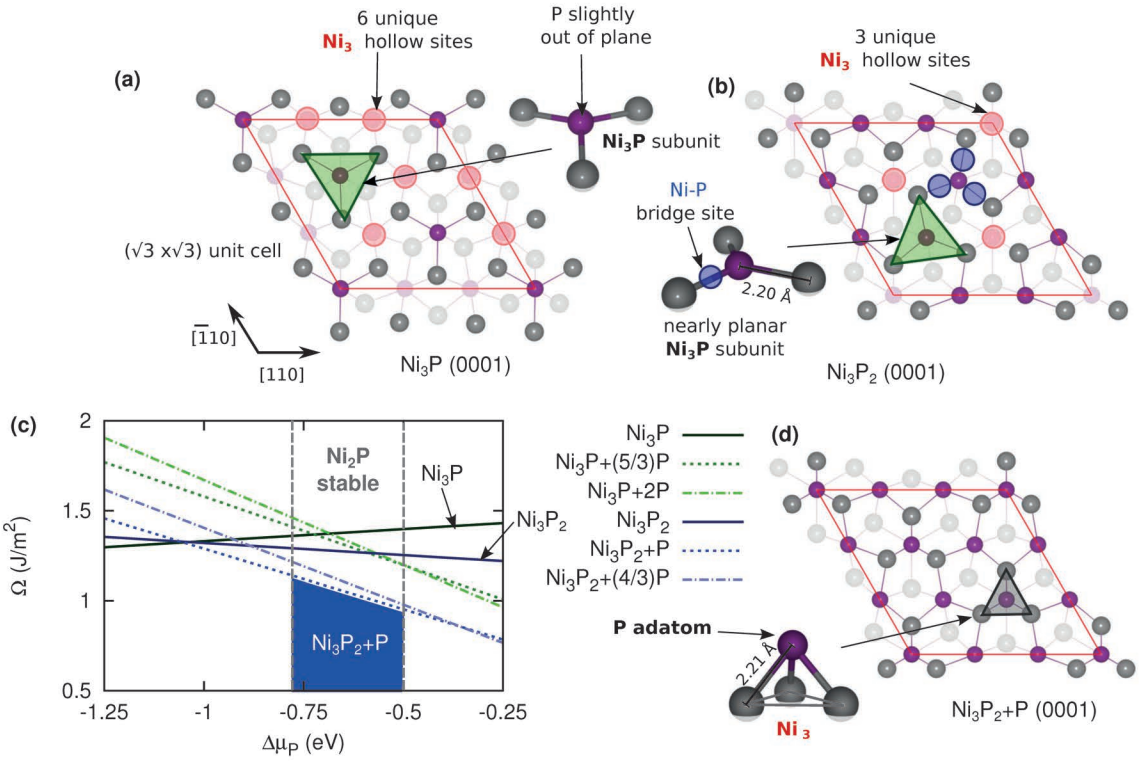

Figure 3.2: Surface crystal structure of $Ni_2P(0001)$ with either a (a) Ni_3P or (b) Ni_3P_2 termination. Only the atoms of the outermost layers are clearly shown. Structural insets highlight the basic subunits of each surface. Red lines outline the $\sqrt{3} \times \sqrt{3}$ $R30°$ supercells. (c) Surface phase diagram for $Ni_2P(0001)$ as a function of $\Delta\mu_P$ (eV). Surface energies are reported in J/m^2. Dashed gray vertical lines border the bulk stability region for Ni_2P.

3.3.3. Surface Structure and Stability of $Ni_5P_4(0001)$ and $(000\bar{1})$

As mentioned above, Ni_5P_4 lacks mirror symmetry along [0001], giving rise to structurally distinct (0001) and $(000\bar{1})$ surfaces. Additionally, the structural layers of Ni_5P_4 are less clearly delineated. We choose to decompose the structure into three layers, and refer to them according to their formula unit: Ni_3P_2, Ni_3P_3, and Ni_4P_3 (Fig. 3.3a). A set of three layers forms half a unit cell, where the addition of a second set of three layers that are translated by half the in-plane lattice vectors completes the full unit cell. Note that due to the absence of a reflection symmetry about the [0001], the layer that comes before a given layer along the [0001] becomes the layer that comes after along the $[000\bar{1}]$. For example the layer below the Ni_5P_4-$Ni_3P_2(0001)$ surface is the Ni_5P_4-Ni_3P_3 layer, whereas the layer below the Ni_5P_4-$Ni_3P_2(000\bar{1})$ is the Ni_5P_4-Ni_4P_3 layer. Therefore, the absence of reflection symmetry means different subsurface layers, which will affect the chemistry of the (0001) *vs.* $(000\bar{1})$ surface layer, despite having the same composition.

For brevity, we only discuss here in detail the structural and chemical features of the (0001) surfaces and their reconstructions, while the $(000\bar{1})$ layer that produces the most stable reconstruction is discussed. We discuss each bulk-derived layer and the different structural and compositional perturbations giving rise to their reconstructions. The compositional variations introduced include Ni vacancies (V_{Ni}), P vacancies (V_P), P adatoms ($+x$P, as in the Ni_2P surface reconstructions), and Ni adatoms ($+x$Ni). We also discuss the energetics involved in these reconstructions, revealing the most likely (thermodynamically stable) surfaces of $Ni_5P_4(0001)$.

Ni_3P_2-derived Surfaces of Ni_5P_4

The upper panel of Fig. 3.3b shows the $(\sqrt{3} \times \sqrt{3})R30°$ (3 unit cells) of the bulk-derived Ni_5P_4-Ni_3P_2 layer. This layer is composed of repeated Ni_3P (green shaded triangle) and P subunits. Each isolated P subunit is coordinated to three P atoms found in the layer below. The removal of this lattice P, creating vacancies (V_P), is unfavorable and destabilizes the surface (Fig. 3.4a, compare solid line, Ni_5P_4-Ni_3P_2, and dash-dotted line, Ni_5P_4-Ni_3P_2+V_P). The corners of three adjacent Ni_3P units surround a common point, which creates a Ni_3 hollow site (red shaded circle). Removal of one of these Ni atoms somewhat stabilizes the surface by about 0.25 J/m^2 (Ni_5P_4-Ni_3P_2+V_{Ni}, dotted line), while saturating all the Ni_3 hollow sites with one and up to three P per site (Ni_5P_4-Ni_3P_2+(1,3)P, double dotted and dash-double dotted lines) stabilizes the surface significantly more (≈0.5-0.75 J/m^2).

Ni_3P_3-derived Surfaces of Ni_5P_4

The middle panel of Fig. 3.3b shows the $(\sqrt{3} \times \sqrt{3})R30°$ Ni_5P_4-Ni_3P_3 layer. Closed-chain Ni_3P_3 subunits compose this surface. The Ni_3P_3 subunit is characterized by a triangular P_3 overlaid on a Ni_3 that are rotated 60° relative to each other. The Ni corners of three Ni_3P_3 subunits also converge on a common point which also creates a trinuclear Ni hollow site (red shaded circles). The same is true for the three P corners which correspondingly create a trinuclear P hollow site (purple shaded circles). As in the Ni_5P_4-Ni_3P_2 surface, saturation of these Ni_3 and P_3 sites with one P per site stabilizes this surface. Fig. 3.4b shows the stabilizing effect of saturating both the Ni_3 and P_3 sites (solid, Ni_5P_4-Ni_3P_3 *vs.* dash-double dotted line, Ni_5P_4-Ni_3P_3+2P). However, saturating the Ni_3 sites alone and removing one of the lattice P (Ni_5P_4-Ni_3P_3+V_P+P) better stabilizes the surface, with one of the remaining P displaced

toward the center of the Ni_3P_2 subunit (see Fig. A.2 in Appendix A). The removal of another P from the Ni_3P_3 subunit further lowers the surface energy in the low $\Delta\mu_P$ regime (dotted line, Ni_5P_4-Ni_3P_3+2V_P+P) and exhibits the same displacement of the remaining P to the center of the Ni_3 base, creating a big Ni_3P subunit (Fig. 3.3c, left panel). The $n$$V_P$+P where (n=1-2) stabilizes the surface by ≈0.25 J/m^2 making these the most stable reconstructions of the Ni_5P_4-Ni_3P_3 layer. The dominance of these surfaces is made possible by a combination of low P chemical potential and the tendency of the system to maximize the number of Ni-P bonds.

Ni_4P_3-derived Surfaces of Ni_5P_4

The bottom panel of Fig. 3.3b shows the $(\sqrt{3} \times \sqrt{3})R30°$ Ni_5P_4-Ni_4P_3 layer. This layer is composed of Ni_4P_3 subunits with the same motif as the Ni_3P_3 subunits found in the Ni_5P_4-Ni_3P_3 layer, but with an additional Ni at the center. Also, similar to the Ni_5P_4-Ni_3P_3 layer, the subunits create Ni_3 (red shaded circles) and P_3 hollow (purple shaded circles) sites (one of each per 1×1 surface). However the extra Ni, which makes this layer distinct from the Ni_5P_4-Ni_3P_3, gives rise to Ni-Ni bridge sites (blue shaded circle). Both Ni vacancies (V_{Ni}) and Ni adatoms (+Ni), marginally stabilize the surface. Saturation with P adatoms of all the Ni_3 and P_3 sites and one Ni-Ni bridge site per Ni_4P_3 subunit (Ni_5P_4-Ni_4P_3+3P) proved to be the most thermodynamically favorable. Partial occupation of the Ni-Ni bridge site may also be achieved, *e.g.* in the case of Ni_5P_4-Ni_4P_3+(8/3)P. The structure of P-saturated Ni_5P_4-Ni_4P_3 (Ni_5P_4-Ni_4P_3+3P) is shown in the right panel of 3.3c. A distinctive trigonal-pyramidal P_4 subunit forms from the adsorption of P at the P_3 site, while a P_2 moiety is stabilized from the bonding of the two P adatoms on a Ni-Ni bridge and a Ni_3 sites.

Global stability of the (0001) Reconstructions of Ni_5P_4

Bulk Ni_5P_4 was calculated to be stable between $\Delta\mu_P$ -0.5 to -0.28 eV, so we examine the surfaces with the lowest surface energy within this range. The Ni_5P_4-Ni_3P_3-derived surface with either the $2V_P$+P and V_P+P reconstruction is found to be the most stable (0001) termination from $\Delta\mu_P$ -0.5 to -0.41 eV (blue shaded region in Fig. 3.4b). At a higher $\Delta\mu_P$, -0.41 to -0.28 eV, the P-covered Ni_5P_4-Ni_4P_3-derived surfaces, Ni_5P_4-Ni_4P_3+3P and Ni_5P_4-Ni_4P_3+(8/3)P, dominate (green shaded region in Fig. 3.4c).

$Ni_5P_4(000\bar{1})$ Reconstructions

Despite the change in the sublayer composition supporting each termination (due to the change of the layer sequence when going the opposite direction) the $(000\bar{1})$ surface is found to be also thermodynamically dominated by the P-covered Ni_5P_4-Ni_4P_3-derived surfaces. The Ni_5P_4-$Ni_4P_3(000\bar{1})$ exhibits the same sites as the Ni_5P_4-$Ni_4P_3(0001)$; *i.e.* Ni_3 and P_3 hollow, and Ni-Ni bridge sites; and similarly composed of the Ni_4P_3 subunit (Fig. 3.3d, left panel). The difference between the (0001) and $(000\bar{1})$ surfaces is the direction of the corrugation of central Ni of the subunits. The central Ni points into the slab in the case of the (0001), while it points outward to the vacuum in the case of the $(000\bar{1})$. Full saturation of the hollow sites and one Ni-Ni bridge site per Ni_4P_3 subunit (Ni_5P_4-Ni_4P_3+3P) also ultimately stabilizes this surface, in fact dominating the whole stability region of Ni_5P_4 (Fig. 3.4d). The compositional and structural similarity of the most stable reconstructions of these (0001) and $(000\bar{1})$ surface isomers may render the two surfaces indistinguishable without the knowledge of the stacking direction within the bulk. However, at slightly lower P concentrations ($\Delta\mu_P <$ -0.41 eV), the (0001) undergoes a phase change into Ni_5P_4-Ni_3P_3+$2V_P$+P,

while the $(000\bar{1})$ robustly remains as Ni_5P_4-Ni_4P_3+3P. See Appendix A for a complete survey of the $(000\bar{1})$ layers and reconstructions.

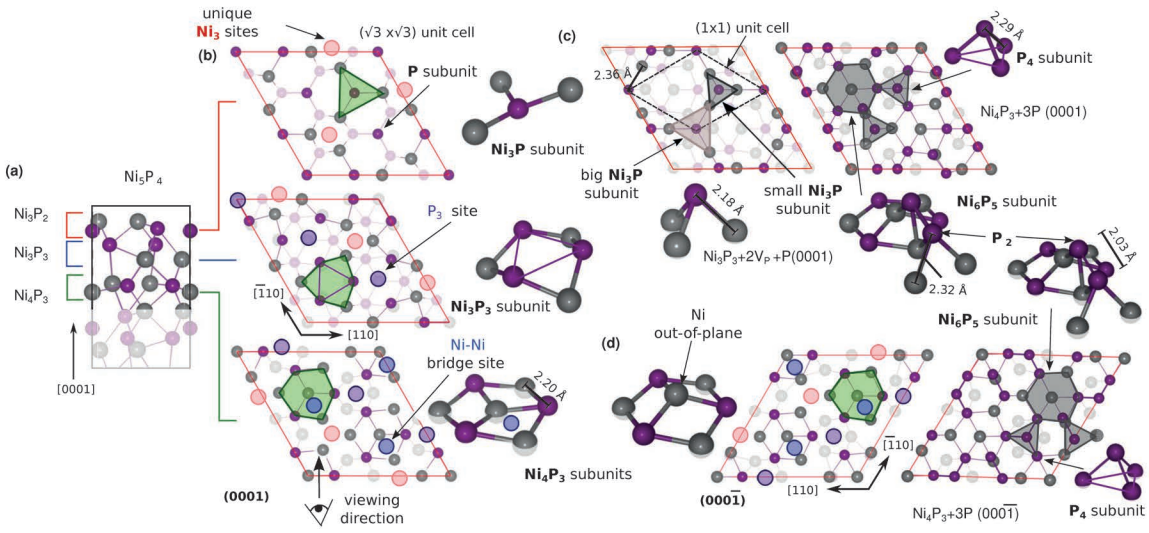

Figure 3.3: Surface crystal structure for bulk-derived terminations and reconstructions of $Ni_5P_4(0001)$ and $(000\bar{1})$. (a) Bulk layering in Ni_5P_4. (b) Bulk-like (0001) terminations Ni_3P_2 (top), Ni_3P_3 (middle), and Ni_4P_3 (bottom) with shaded regions corresponding to the insets highlighting important structural features. (c) Stable (0001) reconstructions Ni_5P_4-Ni_3P_3+2V_P+P (left) and Ni_5P_4-Ni_4P_3+3P (right). (d) Bulk-like (left) and reconstructed (right) Ni_5P_4-Ni_4P_3+3P$(000\bar{1})$.

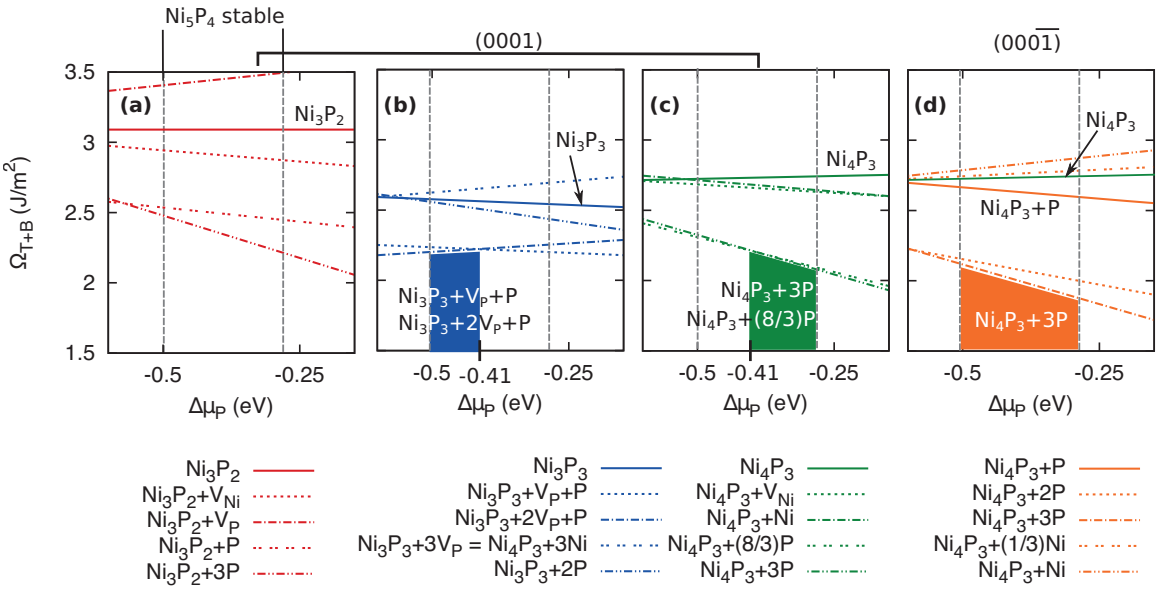

Figure 3.4: Surface phase diagram for $Ni_5P_4(0001)$ and $(000\bar{1})$ surfaces as a function of $\Delta\mu_P$ (eV). Ω_{T+B} corresponds to the combined surface energies (J/m^2) of the top (T) and bottom (B) surfaces of (a) $Ni_3P_2(0001)$, (b) $Ni_3P_3(0001)$, (c) $Ni_4P_3(0001)$, and (d) $Ni_4P_3(000\bar{1})$ and their respective reconstructions. Shaded areas denote regions of $\Delta\mu_P$ where certain reconstructions (as labeled) are favored.

3.3.4. Comparison of the Stable Ni_2P and Ni_5P_4 (0001) Reconstructions

We have shown that the enrichment of Ni_2P and Ni_5P_4 (0001) and (000$\bar{1}$) surfaces with P has a significant stabilizing effect. The most stable reconstruction for Ni_2P(0001) is the Ni_3P_2+P surface, which has an overall composition of Ni_3P_3 and a 1:1 ratio of Ni to P. The preferred reconstructions for Ni_5P_4 (0001) and (000$\bar{1}$), namely Ni_3P_3+[1-2]V_P+P and Ni_4P_3+[(8/3)-3]P, where Ni/P varies from 0.66 to 1.5. The reconstructions of Ni_2P and Ni_5P_4 similarly exhibit P adsorption at Ni_3 hollows to generate a Ni_3P trigonal pyramid, the only difference being the number of Ni_3 hollow sites per unit area and the Ni-P bond distance (see Figs. 3.2 and 3.3). Unlike Ni_2P reconstructions, however, Ni_5P_4-Ni_4P_3+3P (0001) and (000$\bar{1}$) promote the formation of stable, aggregates of P on their surfaces, such as P_4 subunits and P_2 moieties. The former closely resembles the tetrahedral P structures in P (*s*, white), has a very similar P-P bond distance ($d_{\mathrm{P_{White}}}$ = 2.21 Å (187) *vs.* d_{Surf} = 2.29 Å) and carries a nearly neutral charge (0.07*e*, Table 3.1), suggesting the possible nucleation of a secondary pure P phase. The P_2 moiety (d_{Surf} = 2.03 Å) has a similar bond length to gaseous P_2 ($d_{\mathrm{P_2}(g)}$ = 1.89 Å) (188) and also posses nearly neutral charge (-0.05*e*, Table 3.1).

From the dominant structures, we see that the major driving force for the P-rich reconstruction is to saturate Ni coordination where it either achieve a tetrahedral or square pyramidal coordination as in the bulk. The complexity of the Ni_5P_4 reconstructions, especially the Ni_4P_3+3P surface composition, presents other avenues for surface stabilization which include the expansion of surface P valency through bonding with additional P, *e.g.* at P_3 hollow sites. P-P bonds may act as possible nucleation centers for white P, $P_2(g)$, and other P phases depending on the environmental conditions.

Unlike transition metal oxides which are more strongly ionic and whose surface reconstructions are governed by surface charge passivation (50; 58; 189), the metallic nickel phosphides are demonstrated to be primarily driven by saturation of the coordination chemistry of the surface atoms. Due to the much greater electronegativity of O compared to P, transition metal oxides also tend to have larger bond dipoles than Ni_2P and Ni_5P_4. Thus, at oxide surfaces, dangling bonds lead to the buildup of a large surface dipole, which can be screened by reconstruction with the constituent elements of the material or by cations or anions readily available in the environment. (49; 190; 191) The geometry of oxide and phosphide reconstructions is also affected by their electron configurations. Since O has a lower valence than P, oxide reconstructions tend to involve O adatoms bridging between two metal atoms or forming coordinatively unsaturated sites (57)), whereas P may easily form three bonds using its $3p$ orbitals while also expand its valence shell using its energy-accessible $3d$ orbitals. On this basis, P can introduce modes of coordination beyond sp bonding. This is clearly demonstrated by the partial population of the P $3d$ orbitals (see Table 3.1 for the $3d$ charge population) and their hybridization with the P $3p$ states (see Fig. A.5 in Appendix A for the projected density of states of some of the surfaces). Such flexibility in valence enables the formation of exotic surface structures including small covalent P clusters.

Because bulk-like terminations of Ni_2P and Ni_5P_4 tend to reconstruct, the previously proposed mechanism of HER (33; 183) on Ni_2P and Ni_5P_4, as well as other chemical reactions (165; 184–186) for which it is catalytically active, must be reevaluated. Depending on the chemical reaction catalyzed, the catalyst may be subsequently exposed to different solvents, ions, molecules, and even applied potentials during operation which would then facilitate further evolution of the surfaces. Here, we provide

a set of chemically sound structures that lay the foundation for studying the effects of these additional variables on catalysis. A few of the most notable routes for additional reconstructions to consider are perhaps through hydration, protonation, and hydroxylation of surface Ni and P species once exposed to water. We show however that the non-stoichiometric P ad-species prevalent on these reconstructions already constitute strong Ni-bond-passivating components of the surface. This suggests that P redox and acid-base chemistries (instead of nickels) are more relevant in determining the eventual transformation and activity of the surfaces during catalysis. Thus, for example phosphine and phosphorous/phosphoric acid chemistries may very well hold the key in understanding Ni-P catalysis in an aqueous environment.

Table 3.1: Calculated Löwdin charges for the surface P atoms for some of the stable reconstructions of Ni_2P and Ni_5P_4.

Composition		**Location**	**No. Electrons**			**Net Charge**
Bulk	*Surface*		$3s$	$3p$	$3d$	$[e]$
Ni_2P	Ni_3P_2+P	Ni_3P subunit	1.60	3.19	0.88	-0.68
		Ni_3-hollow	1.75	3.09	0.48	-0.30
Ni_5P_4 (0001)	Ni_4P_3+3P	Ni_4P_3 subunit	1.46	3.06	1.09	-0.61
		Ni_3-hollow	1.55	3.05	0.72	-0.32
		P_3-hollow	1.55	2.77	0.61	0.07
		Ni-Ni bridge	1.63	2.93	0.49	-0.05

3.4. Conclusions

Both Ni_2P and Ni_5P_4 favor P-rich reconstructions on the (0001) facet, arising from the extended coordination of P and the tendency of Ni to remain saturated. For Ni_2P, the most stable surface termination is Ni_3P_2+P, where P adatoms rest on the Ni_3-hollow sites. On $Ni_5P_4(0001)$, reconstructions of both the Ni_3P_3 and Ni_4P_3 surface layer compositions may form. For its $(000\bar{1})$, Ni_5P_4-Ni_4P_3+3P is most stable. Ni_5P_4-Ni_4P_3+3P(0001) and $(000\bar{1})$ are furnished by stable P clusters, P_4 and P_2, which resemble phases of P, namely white P and $P_2(g)$ respectively. The ability of P to agglomerate on $Ni_5P_4(0001)$ sets this phase apart from $Ni_2P(0001)$ and, more generally, from transition metal oxides, which are much more limited in their modes of reconstruction and more adherent to surface charge passivation requirements.

CHAPTER 4 : Mechanism of H_2 evolution on aqueous reconstructions of the (0001) surface of Ni_2P and Ni_5P_4: the crucial role of phosphorus

4.1. Introduction

Electrical energy produced from renewable sources, *e.g.* solar cells, can be used to split water to produce fuel (H_2). The cathode half-reaction of water splitting is the hydrogen evolution reaction (HER), which can be performed in both acidic (shown below) and to some extent in basic aqueous media:

$$2H^+ + 2e^- \rightleftharpoons H_2(g). \tag{4.1}$$

Pt is currently considered as the benchmark catalyst for the HER; however, it is both scarce and expensive. (192) This has motivated many scientists in the last decade to search for earth-abundant, cheap alternatives to Pt as electrocatalysts for the HER. (193; 194)

There are two well-known mechanisms for the HER: Volmer-Tafel (VT) and Volmer-Heyrovsky (VH). (195–201) Their chemical representations are as follows:

$$H^+ + S + e^- \rightleftharpoons S - H \text{ (Volmer step)} \tag{4.2}$$

$$S - H + S' - H \rightleftharpoons S/S' + S'/S - H_2 \text{ (Tafel step)} \tag{4.3}$$

$$S - H + H^+ + e^- \rightleftharpoons S - H_2 \text{ (Heyrovsky step)} \tag{4.4}$$

$$S - H_2 \rightleftharpoons S + H_2(g) \tag{4.5}$$

Both mechanisms start with the Volmer step, where one H binds to a site (S) on the electrocatalyst surface. From here, the reaction can proceed in two different ways. Either another H will bind at a separate site (Tafel step), where S and S′ may or may not be of the same type, or directly on top of an adsorbed H forming an H_2 complex (Heyrovsky step). Both the Tafel and Heyrovsky steps are followed by the desorption of $H_2(g)$ or alternatively, lead to direct desorption of the molecule.

A few potential substitutes for Pt HER electrocatalysts are molybdenum sulfides (193; 202–215), molybdenum and tungsten carbides, (216–226) nitrides, (227–233) and nickel phosphides. (32–34; 36; 183; 184; 234) The hydrogen evolution activity of $Ni_2P(0001)$ was originally predicted computationally (33) and then subsequently demonstrated experimentally. (32) It was proposed that the presence of P deactivates Ni and decreases the number of metal-hollow sites (the so-called "ensemble effect") while also providing weak binding for H at Ni-P bridge sites. (184) The electrostatic attraction between adsorbed $H^{\delta+}$ (at Ni-P bridge sites) and $H^{\delta-}$ (at trinuclear Ni_3-hollow sites) was proposed to facilitate HER on Ni_2P. (32)

Surface phase stability is of the utmost importance in predicting the performance of heterogeneous catalysts. (49–51; 58; 67; 189–191; 235–237) Different surface structures give rise to different sites for H adsorption, which may lead to different HER mechanisms and overpotentials. Therefore, it is essential to determine the structure and composition of the catalyst surface(s) under fabrication and operating conditions. Ni_2P has two bulk layers, Ni_3P and Ni_3P_2, along the crystal's (0001) direction; accordingly there are two bulk-like terminations: Ni_3P and Ni_3P_2. DFT calculations predict that the latter is more stable under Ni_2P bulk formation conditions (172). Scanning tunneling microscopy (STM) and dynamic low-energy electron diffraction (LEED) experiments reveal that nonstoichiometric additional P covers $\approx 80\%$ of this surface;

hereafter, the surface found experimentally is denoted as $Ni_2P(s)/Ni_3P_2(0001)$+P (note that the naming scheme we adopt here is bulk/surface+adlayer where each term is the empirical formula). (171; 173–175; 182) Calculated adsorption energies for H at low coverage on different sites of $Ni_2P(s)/Ni_3P_2(0001)$+P show that H prefers to bind on these P adatoms. (183) We recently studied the reconstructions of both $Ni_2P(0001)$ and $Ni_5P_4(0001)/(000\bar{1})$ surfaces with DFT and discovered that the most stable terminations, subject to an inert environment, are P-enriched. (67) We predict (67) that $Ni_2P(0001)$ prefers a P-covered reconstruction of the Ni_3P_2 termination that is consistent with STM and LEED experiments in the literature. (171; 173–175; 182) Ni_5P_4 has three bulk layers along the (0001) and $(000\bar{1})$ directions: Ni_4P_3, Ni_3P_3, and Ni_3P_2. $Ni_5P_4(0001)/(000\bar{1})$ favors P-enrichment of the Ni_4P_3 and Ni_3P_3 terminations. Ni_3- and P_3-hollow sites, the latter of which are only present on $Ni_5P_4(s)/Ni_4P_3(0001)/(000\bar{1})$, bind additional P.

Recently, Ni_5P_4 was synthesized and was shown to exhibit exceptional, Pt-level performance for HER at $p\mathrm{H} \approx 0$ and applied potentials ranging from $U = 0$ V to -0.1 V *vs.* the standard hydrogen electrode (SHE). (34) The superior performance of Ni_5P_4 was attributed to a higher positive charge on Ni atoms and to the ensemble effect of P, where the number of Ni_3-hollow sites that bind H very strongly is decreased due to the abundance of P, which therefore leads to more thermoneutral adsorption. (33; 184; 234) Additionally, it was shown that monodisperse Ni_5P_4 nanocrystals have higher surface area and greater stability in acidic media than Ni_2P. In this chapter, we only consider Ni_2P and Ni_5P_4 because they are experimentally demonstrated to be the most HER-active nickel phosphide polymorphs. (234; 238; 239) Additionally, these two compounds do not have strong differences between their crystal structures and structural motifs in the bulk, both compounds also have comparable electronic

conductivities. (167–171) Therefore, conclusions on their relative HER activities based on the compositional and structural properties of their surfaces, as well as the relative stabilities of their bulk and surfaces, can be made. A better understanding of the atomistic mechanism of the HER for various Ni phosphide systems will accelerate the design and fabrication of robust HER electrocatalysts.

To address this need, we apply DFT calculations and thermodynamics to predict the most stable surface phases under normal synthetic conditions and in an electrochemical environment, *i.e.* aqueous solution at specified *p*H and applied potential U (hereafter U is implicitly relative to SHE). Each of the phosphides we model has been electrochemically investigated in aqueous solution (32; 34) and is found to be stable. Thus, we model the catalytic properties of each bulk phase and evaluate which surfaces are responsible for their respective activities. We calculate the free energy of reaction of the elementary steps involved in the electrochemical HER via first principles, and consequently reveal the lowest-energy pathways on $Ni_2P(0001)$ and $Ni_5P_4(0001)/(000\bar{1})$.

4.2. Methods

4.2.1. First Principles Calculations

DFT (72; 177) calculations were carried out using the Quantum ESPRESSO (version 5.1) software. (240) Geometric relaxation of the bulk and surface structures was performed until changes in the total energy and force were less than 1.4×10^{-3} eV/cell and 2.6×10^{-2} eV/Å respectively. Optimized, (106) norm-conserving, designed nonlocal (105) pseudopotentials were constructed using the OPIUM (version 3.7) software (179) to replace the core electrons and nucleus with a smoother, effective potential. We used the generalized gradient approximation (GGA) as formulated by

Perdew, Burke, and Ernzerhof (PBE) to calculate electron exchange and correlation energies. (87) The valence orbital wavefunctions were expanded in a plane-wave basis with a cutoff energy of 680 eV. Gaussian electronic smearing of 0.07 eV was applied to the band occupations near the Fermi energy to improve electronic k-point convergence. We used the semiempirical DFT-D2 method (129) to include van der Waals (vdW) interactions, which are vital for modeling catalytic transformations. (241; 242)

Bulk lattice constants were also relaxed with a pressure convergence threshold of 6.3×10^{-6} eV/Å^3. The total energies of bulk Ni_2P and Ni_5P_4 were found to be converged with 5×5×6 and 5×5×4 grids respectively, offset along k_z. Calculated lattice constants and formation energies of Ni(s), P(s,white), Ni_2P(s), and Ni_5P_4(s) are found to be in good agreement with experimental values. (67) Slab models for $Ni_2P(0001)$ and $Ni_5P_4(0001)/(000\bar{1})$ were generated with $\sqrt{3}\times\sqrt{3}\ R30°$ surface unit cells and ≈ 25 Å of vacuum space separating layers. Accordingly, the k-point grid was reduced to 3×3×1. A dipole correction was added to the center of the vacuum region to cancel any artificial electric fields between the slabs. (180) Vibrational frequencies of adsorbates and surface atoms directly coupled to them were calculated (from truncated Hessian matrices) using density functional perturbation theory (DFPT). The charge densities used for these calculations were obtained by lowering the total energy convergence threshold for SCF calculations from 1.4×10^{-5} eV/cell for geometry relaxations to 1.4×10^{-9} eV/cell.

4.2.2. Theory

In order to accurately model catalysis, it is necessary to construct a realistic model of the surface under experimental conditions, *i.e.* in aqueous solution with electrochemical driving forces, pH and U. (54) This can be achieved by considering the equilibrium between surface atoms and adsorbates, and their aqueous counterparts.

When a surface is in contact with an aqueous solution, the surface can lose atoms to or gain atoms from the solution. The chemical equation defining the equilibrium between a surface and aqueous solution where atom A is being exchanged is as follows

$$\mathrm{S-A}+n_w\mathrm{H_2O}\rightleftharpoons\mathrm{S}+[\mathrm{H}_x\mathrm{AO}_y]^z+n_\mathrm{H}\mathrm{H}^+ + n_e e^- \tag{4.6}$$

where S is the surface, n_w is the number of water molecules needed to oxidize (reduce) and/or solvate A in the solution, $[\mathrm{H}_x\mathrm{AO}_y]^z$ is the most stable aqueous phase of A, z is the charge of the A-complex which can either be positive, negative, or zero, n_H is the number of protons formed, and n_e is the number of electrons released. Note that the latter two can be negative in which case proton(s) and electron(s) are gained to form $[\mathrm{H}_x\mathrm{AO}_y]^z$. The free energy change for this dissolution reaction of one A ion is

$$\Delta G_\mathrm{A,diss}=(G_\mathrm{S}-G_\mathrm{SA})+(G_{\mathrm{H}_x\mathrm{AO}_y^z}+n_\mathrm{H}G_\mathrm{H}+n_eG_e-n_wG_{\mathrm{H_2O}}) \tag{4.7}$$

We can rewrite $\Delta G_\mathrm{A,diss}$ with respect to the standard state of A, A(std), as

$$\Delta G_\mathrm{A,diss}=(G_\mathrm{S}+G_\mathrm{A(std)}-G_\mathrm{SA})+(G_{[\mathrm{H}_x\mathrm{AO}_y]^z}+n_\mathrm{H}G_\mathrm{H}+n_eG_e-G_\mathrm{A(std)}-n_wG_{\mathrm{H_2O}}) \tag{4.8}$$

Note here we simply add $G_\mathrm{A(std)}$ to the first term and subtract it from the second term. We define the first term as the differential desorption (dsrp) free energy ΔG_dsrp for A to leave the surface and form A(std). We re-express the second term with respect to the standard oxidation/reduction free energy $\Delta G^\circ_{\mathrm{A(std)}/\mathrm{H}_x\mathrm{AO}_y^z}$ of A(std) to form $[\mathrm{H}_x\mathrm{AO}_y]^z$ (see additional theoretical details and Table B.1 in Appendix B). Our final expression is

$$\Delta G_\mathrm{A,diss}=\Delta G_\mathrm{dsrp}+\Delta G^\circ_{\mathrm{A(std)}/\mathrm{H}_x\mathrm{AO}_y^z}+k_\mathrm{B}T\ln a_{\mathrm{H}_x\mathrm{AO}_y^z}-2.303n_\mathrm{H}k_\mathrm{B}T p\mathrm{H}-n_eq_eU \tag{4.9}$$

The last three terms represent the deviation of the chemical potentials due to concentrations of the aqueous species: $H_xAO_y^z$, and H^+ from the standard condition, and of the electronic energy relative to SHE:

$$\Delta G_{\mathrm{A,diss}} = \Delta G_{\mathrm{dsrp}} + \Delta G^{\circ}_{\mathrm{A(std)}/H_xAO_y^z} + \Delta\mu_{H_xAO_y^z} + n_{\mathrm{H}}\Delta\mu_{\mathrm{H}} + n_e\Delta\mu_e \tag{4.10}$$

where $\Delta\mu_{H_xAO_y^z} = k_{\mathrm{B}}T \ln a_{H_xAO_y^z}$, $\Delta\mu_{\mathrm{H}} = -2.303 k_{\mathrm{B}}T p\mathrm{H}$, and $\Delta\mu_e = -q_eU$. These terms can be thought of as knobs that control the chemical potentials of $H_xAO_y^z$, protons, and electrons respectively. We calculated ΔG_{dsrp} using DFT and evaluated the relative stability of different surface phases of Ni_2P and $Ni_5P_4(0001)$ by modeling reactions represented in Eq. 4.6. We investigated various H coverages on each surface reconstruction, up to 7 H atoms per 3 surface unit cells ($\theta = 7/3$ or 1.32 nmol H/cm^2, where the surface area, A = 2.93×10^{-15} cm^2/(1×1) surface) for Ni_2P and $\theta = 14/3$ H atoms (2.00 nmol H/cm^2, where A = 3.87×10^{-15} cm^2/(1×1) surface) for Ni_5P_4, in increments of $\Delta\theta = 1/3$ (0.19 nmol H/cm^2 for Ni_2P and 0.14 nmol H/cm^2 for Ni_5P_4). The aforementioned maximum coverages are the highest possible coverages for which H adsorption is preferred to physisorbed H_2. We found that the thermodynamically optimal coverages in the relevant range of pH and applied potential are within the maximum coverages explored. We computed their free energies using Eq. 4.10 relative to the Ni_3P_2+P and bulk Ni_4P_3 terminations for Ni_2P and Ni_5P_4 surfaces, respectively (see Tables B.2-B.4 in Appendix B).

We generate Pourbaix diagrams, which map the equilibrium phases of an aqueous electrochemical system, for Ni and P, by calculating $\Delta G^{\circ}_{\mathrm{A(std)}/H_xAO_y^z}$ and identifying the most stable species for given pH and U (see Fig. B.1 in Appendix B). An expression for the complete dissolution of the bulk Ni_aP_b phase can be derived following

Eqs. 4.6-4.10:

$$\Delta G_{\mathrm{Ni}_a\mathrm{P}_b,\mathrm{diss}} = -\Delta G^f_{\mathrm{Ni}_a\mathrm{P}_b} + a\Delta G_{\mathrm{Ni}(s)/\mathrm{H}_x\mathrm{NiO}_y^z} + b\Delta G_{\mathrm{P}(s,\mathrm{white})/\mathrm{H}_x\mathrm{PO}_x^z} \tag{4.11}$$

where $\Delta G^f_{\mathrm{Ni}_a\mathrm{P}_b}$ is the free energy of formation of $\mathrm{Ni}_a\mathrm{P}_b$. The stability criterion is $\Delta G_{\mathrm{Ni}_a\mathrm{P}_b,\mathrm{diss}} \geq 0$, thus we define the bulk stability boundary as

$$\Delta G^f_{\mathrm{Ni}_a\mathrm{P}_b} \leq a\Delta G_{\mathrm{Ni}(s)/\mathrm{H}_x\mathrm{NiO}_y^z} + b\Delta G_{\mathrm{P}(s,\mathrm{white})/\mathrm{H}_x\mathrm{PO}_y^z} \tag{4.12}$$

The bulk phase diagrams of Ni_2P and Ni_5P_4 are shown in Fig. B.2 in Appendix B. Finally, we calculated free energies of hydrogen adsorption (see reaction in Eq. 4.2)

$$\Delta G_{\mathrm{ads}} = G(n_{\mathrm{H}} + 1) - G(n_{\mathrm{H}}) - G(\mathrm{H}^+ + e^-) \tag{4.13}$$

at 298.15 K, pH = 0, and U = 0 V.

4.3. Results and Discussion

4.3.1. Structure and Aqueous Stability of Ni_2P(0001) Surfaces

For this study, we consider an acidic aqueous environment (pH = -0.1–1, which corresponds to typical experimental conditions for HER in acid, *i.e.* $[H_2SO_4]$ = 1–0.1 M) (32; 34) and calculate the free energy of Ni_2P(0001) surfaces in equilibrium with 1 M Ni^{2+} or Ni(s), and 1 M PH_3 or 1 M H_3PO_4 at 298.15 K (see Table B.2 in Appendix B). For the purpose of our discussion, we choose the standard concentration of 1 M for the aqueous species as a suitable reference to define the bulk and surface phase stability boundaries. The results for 0.01 and 0.001 M concentrations are plotted in Fig. B.6 in Appendix B, where we show that qualitatively the observed trends in

stability and reactivity are unaffected, while quantitatively, the upper bound for the applied potential where the phases are stable varies by 0.02–0.05 V. In the regions of U and pH where solubility is greater than 1M, it is reasonable to conclude that the catalyst would no longer be practical, as it would experience significant material loss during catalysis and after prolonged use (see Fig. B.2 for the calculated solubility of the bulk as a function of U and pH). Also, we note that the overpotential calculated for an HER elementary step for a given surface is independent of these concentrations. Fig. 4.1A shows the phase diagram for the (0001) surface of Ni_2P. At $U >$ -0.21 V, $Ni_2P(s)/Ni_3P_2(0001)$ with one H per surface unit cell, hereafter denoted as $Ni_2P(s)/Ni_3P_2(0001)$+H, is the dominant surface phase. The $\sqrt{3} \times \sqrt{3}$ $R30°$ surface has one H bonded to each Ni_3-hollow site (labeled 1-3 in Fig. 4.1B). Ni_3-hollow sites strongly bind H, with an adsorption free energy (see Eq. 4.13) of -0.47 eV/H. We summarize ΔG_{ads} for different sites on $Ni_2P(0001)$ and $Ni_5P_4(000\bar{1})$ in Table 4.1.

For -0.21 V $\geq U \geq$ -0.36 V, a P-enriched phase, $Ni_2P(s)/Ni_3P_2(0001)$+P+(7/3)H is most stable. Aqueous P reacts with the surface, replacing the H atom at each Ni_3-hollow site with P. This surface is similar to the UHV reconstruction $Ni_2P(s)/Ni_3P_2(0001)$+P, (67) but here P-adatoms are also hydrogenated. More specifically, one adatom P forms a surface PH_3 moiety, whereas the other two form PH_2 for a total of seven H atoms (labeled 1-7 in 4.1C) per $\sqrt{3} \times \sqrt{3}$ $R30°$ surface unit cell. These units, with an average P-H bond length of 1.43 Å, form the precursor for phosphine molecule ($\bar{d}_{\text{PH}} = 1.42$ Å in Ref. 243) desorption, which occurs at $U <$ -0.36 V. The side view in Fig. 4.1B reveals that P is exposed and able to react with more H than the $Ni_2P(s)/Ni_3P_2(0001)$ surface. The bond length between in-plane Ni and P is 2.19 Å, whereas the bond length between Ni and adatom P is elongated (2.35 Å), signaling a weaker bond. As such, adatom P is able to form a

stronger bond with H ($\Delta G_{\text{ads}} = 0.05$ eV/H) than the lattice P ($\Delta G_{\text{ads}} = 0.27$ eV/H). H-binding is much weaker on adatom P sites than Ni_3-hollow sites, in agreement with the literature. (183)

Below U = -0.36 V, PH_3 desorbs and re-exposes the Ni_3-hollow sites again for H to bind, one H per Ni_3. Two additional H adsorb on the surface (labeled 4 and 5 in Fig. 4.1D) forming $Ni_2P(s)/Ni_3P_2(0001)$+(5/3)H, where Ni forms a complex with H_2, pushing the central H to a Ni-bridge site. The average Ni-H bond length involving atomic H decreases from 1.79 Å on $Ni_2P(s)/Ni_3P_2(0001)$+H to 1.67 Å on $Ni_2P(s)/Ni_3P_2(0001)$+(5/3)H because of the reduced coordination of number of H. The H-H bond length (0.82 Å) is only slightly larger than that of $H_2(g)$ (0.74 Å in Ref. 244), highlighting the weak adsorption of this species.

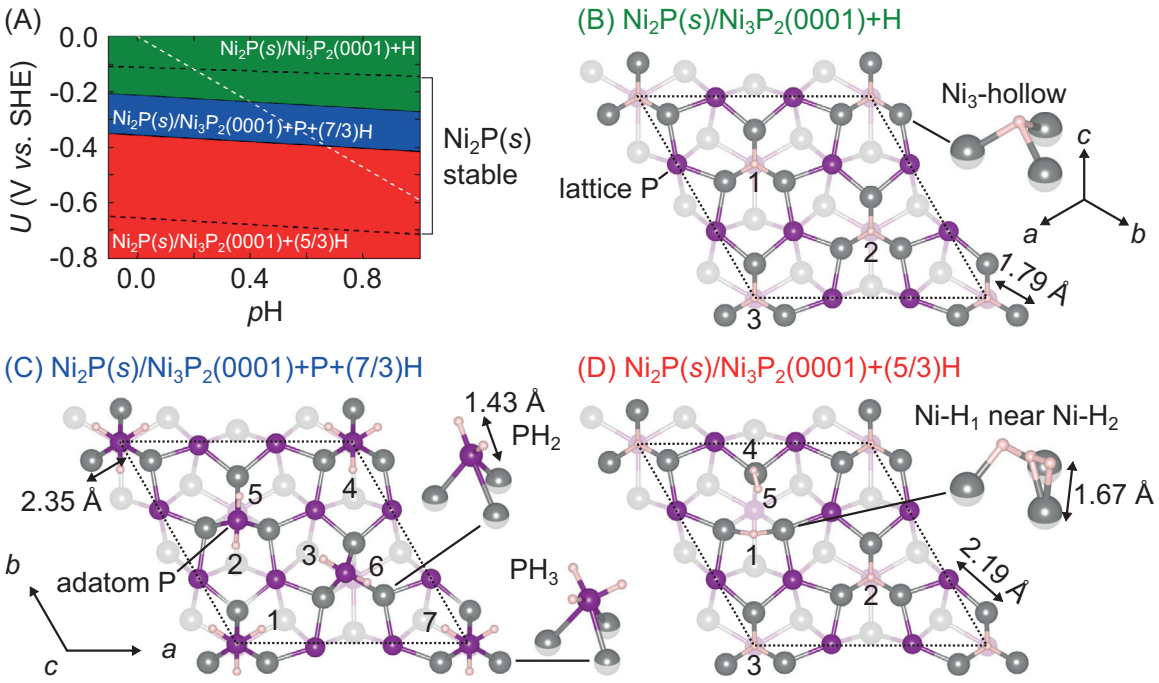

Figure 4.1: (A) Surface phase diagram of $Ni_2P(0001)$ in equilibrium with 1 M Ni^{2+} or Ni(s), and 1 M PH_3 or 1 M H_3PO_4 at 298.15 K. (B)-(D) show the evolution of the surface through adsorption-desorption equilibrium of P. P dissolves off the surface as phosphates (B). As the potential is lowered it redeposits as phosphines (C), up to a point where PH_3 becomes very soluble and re-exposes the Ni sites for H to bind (D). Average bond lengths are indicated.

Table 4.1: Free energy of H adsorption on selected surface sites of $Ni_2P(0001)$ and $Ni_5P_4(000\bar{1})$ in equilibrium with 1 M Ni^{2+} or Ni(s), and 1 M PH_3 or 1 M H_3PO_4 at 298.15 K, $U = 0$ V, and pH = 0.

Bulk	Surface	Orientation	Active Site	ΔG_{ads} (eV)
Ni_2P	Ni_3P_2+H	(0001)	Ni_3-hollow	-0.47
Ni_2P	Ni_3P_2+P+(7/3)H	(0001)	Adatom P	0.01 to 0.14
Ni_5P_4	Ni_4P_3+4H	$(000\bar{1})$	Ni_3-hollow	-0.57
Ni_5P_4	Ni_4P_3+4H	$(000\bar{1})$	P_3-hollow	-0.27 to -0.06

4.3.2. Structure and Aqueous Stability of $Ni_5P_4(000\bar{1})$ Surfaces

For $Ni_5P_4(000\bar{1})$, only two different surface phases are observed under acidic (pH = -0.1–1), reducing (U = 0.0 to -0.8 V) conditions (see Fig. 4.2A). For $U \geq$ -0.37 V, the predominant surface phase is $Ni_5P_4(s)/Ni_4P_3(000\bar{1})$+4H. The structure of this surface is shown in Fig. 4.2B, and it consists of repeating Ni_3- and P_3-hollows connected by central Ni atoms. The P_3-hollow sites are unique to Ni_5P_4 and specifically the Ni_4P_3 termination. There are two different types of Ni-P bonds, one with a bond length of 2.09 Å between P and the central Ni and another with a bond length of 2.26 Å between P and a Ni from a Ni_3-hollow. The average of these two Ni-P bond lengths is near the bond length between in-plane Ni and P on $Ni_2P(s)/Ni_3P_2(0001)$ surfaces. Three H atoms adsorb, per $\sqrt{3} \times \sqrt{3}$ supercell, one at each Ni_3-hollow with equal bonding contributions from all three Ni atoms and an average bond length of 1.74 Å. H binding is stronger at the Ni_3-hollow site on $Ni_5P_4(s)/Ni_4P_3(000\bar{1})$+4H, with an average adsorption free energy of -0.57 eV/H. Nine additional H atoms per supercell adsorb, three per P_3-hollow site (circled and shaded in light purple in in Fig. 4.2B). Each H makes a single P-H bond of length 1.42 Å, and the H atoms point toward the center of the P_3-hollow. We find a positive correlation between ΔG_{ads} and H coverage (n_{H}) at the P_3-hollow sites (see Fig. 4.3) and attribute this to repulsive interactions between H adsorbates. As H coverage at P_3-hollows increases, surface H species are forced into close proximity, thereby increasing the P-P-H angle (see inset in Fig. 4.3), destabilizing H adsorption, and shifting it toward thermoneutrality. This coverage-dependent chemisorption energy is the key feature that makes P_3-hollows superior to other P-based active sites. At more negative potentials, $U <$ -0.37 V, two additional H atoms per supercell bind to one of the Ni_3-hollows, generating a Ni-H_2 complex (d_{HH} = 0.86 Å) identical to that on $Ni_2P(s)/Ni_3P_2(0001)$+(5/3)H, as shown in Fig.

4.2C. Consequently, there are three domains of binding energy on Ni_5P_4 (colored red, blue, and green in Fig. 4.3). The maximum coverage of $Ni_5P_4(000\bar{1})$ is 1.5 times larger than that of $Ni_2P(0001)$. We also consider P-enriched $(000\bar{1})$ surfaces in our stability analysis (see Table B.4 in Appendix B) but find that they are not stable in acidic aqueous media, unlike in $Ni_2P(0001)$, where an adlayer of P forms. P_3-hollows, however, are more advantageous for the HER than adatom P because they exist at less negative overpotentials.

We defer our discussion of $Ni_5P_4(0001)$ surfaces to Appendix B (see Fig. B.3) because the $(000\bar{1})$ surfaces offer lower HER overpotentials, which now will be discussed.

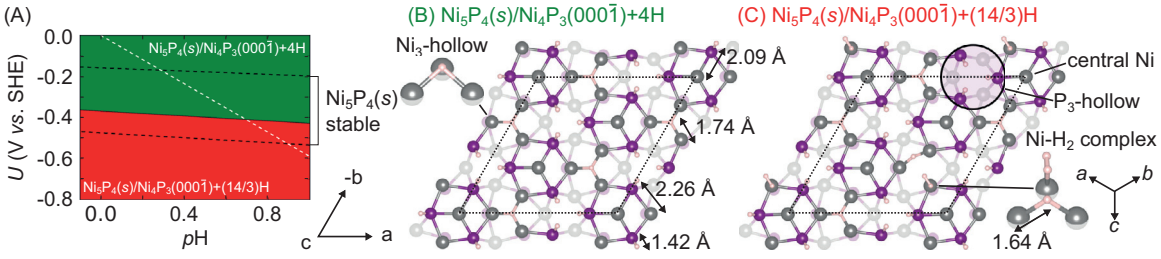

Figure 4.2: (A) Surface phase diagram of $Ni_5P_4(s)/Ni_4P_3(000\bar{1})$ in equilibrium with 1 M Ni^{2+} or $Ni(s)$, and 1 M PH_3 or 1 M H_3PO_4 at 298.15 K. (B)-(C) show the evolution of the surface through adsorption-desorption equilibrium of H. H binds at the Ni_3- and P_3-hollow sites, one H per Ni_3 and three H per P_3. As the potential is lowered, two additional H adsorb at the Ni_3-hollow site, forming a Ni-H_2 complex.

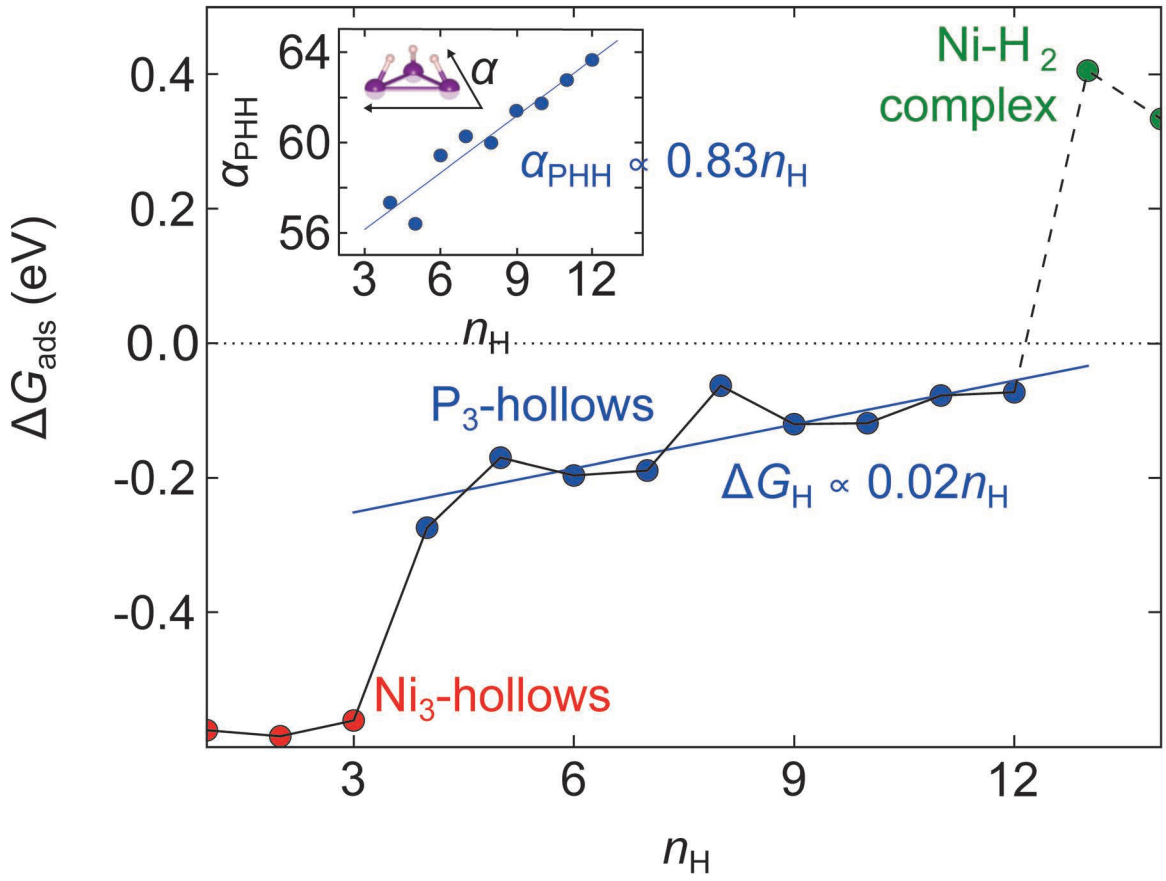

Figure 4.3: Free energy of H adsorption as a function of H coverage (n_{H}) on $Ni_5P_4(s)/Ni_4P_3(000\bar{1})$ in equilibrium with 1 M Ni^{2+} or Ni(s), and 1 M PH_3 or 1 M H_3PO_4 at 298.15 K, $U = 0$ V, and pH $= 0$. Colors differentiate H binding sites. Solid (dashed) lines connect coverages where hydrogen adsorption is exergonic (endergonic). Dotted line at $\Delta G_{\mathrm{H}} = 0$ eV corresponds to thermoneutral H adsorption. We fit ΔG_{H} at P_3-hollow sites to a simple linear model to quantify the destabilization of P-H with increasing n_{H}. Inset is a plot of the P-P-H angle *vs.* n_{H}.

4.3.3. HER Mechanism of $Ni_2P(0001)$ and $Ni_5P_4(000\bar{1})$ Surfaces

In general, the most efficient catalytic mechanisms occur between nearly isoenergetic surface phases, such that each step in the reaction is thermoneutral or low-energy. For electrochemical reactions that do not occur spontaneously, an applied potential can be used to drive the reaction. Here, we define the overpotential η as the potential required to make all elementary steps in the HER spontaneous and ensure catalyst stability, since a prerequisite for catalysis is a regenerable surface that supports it. In the following discussion on the mechanism of the HER, the predictions we make satisfy these criteria.

The HER mechanism on Ni_2P is still debated in the literature. Computational studies have focused on stoichiometric Ni_2P surfaces, with none considering the influence of surface reconstruction driven by the aqueous phase in contact with the catalyst. There are two main proposals for the HER active site on Ni_2P: cooperative Ni_3-hollow and Ni-P bridge sites on the (0001) facet, (33) and Ni-Ni bridge sites on the $(\bar{1}\bar{1}20)$ and $(11\bar{2}0)$ surfaces. (36). Here, we present alternative mechanisms that exhibit the lowest overpotentials for HER on the (0001) surfaces of Ni_2P and Ni_5P_4. A less favorable mechanism on $Ni_5P_4(0001)$ is reported in Appendix B.

For $Ni_2P(0001)$, we find that the H-covered, P-enriched Ni_3P_2+P+(7/3)H surface offers the lowest overpotential for HER via the Volmer-Heyrovsky mechanism. Fig. 4.4A shows the free energies and structures of reaction intermediates. The first step of this reaction involves two concerted events: (a) reaction of a proton in solution and electron with H at a PH_3 moiety and (b) desorption of $H_2(g)$ to form PH_2. In the second step, the PH_3 subunit is replenished by another proton and electron. At $U = 0$ V, this reaction has a 0.14 eV barrier (black line in Fig. 4.4A),

and consequently the application of -0.14 V (red line) makes each step spontaneous. Ni_3P_2+P+(7/3)H, however, is only stable for -0.21 V $\geq U \geq$ -0.36 V. Therefore, we portray an overpotential of -0.21 V (green line) for $Ni_2P(0001)$, since the reaction is limited by the stability of the active surface phase (the dominant surface at $U =$ -0.14 V has a larger overpotential requirement, $U =$ -0.26 V, see Fig. B.4 in Appendix B). This mechanism is different from those previously proposed in the literature for Ni_2P (33; 36), highlighting the importance of considering aqueous phase stability in the prediction of catalytic mechanisms. It was also predicted that Ni-P bridge sites on $Ni_2P(s)/Ni_3P_2(0001)$ offer weaker binding for H. However, we find that these sites are not stable.

$Ni_5P_4(000\bar{1})$ offers a surface phase that provides efficient HER catalysis, Ni_4P_3+4H. Like $Ni_2P(0001)$, this surface favors a Volmer-Heyrovsky mechanism, as shown in Fig. 4.4B. This mechanism involves the simultaneous addition of H and abstraction of $H_2(g)$ from a P_3-hollow site followed by H adsorption to replenish the third H at the P_3-hollow site. At $U = 0$ V, this HER mechanism has a smaller barrier (0.07 eV at $U = 0$ V, black line in 4.4B) and consequently requires a smaller overpotential (-0.07 V, red line) to make each step spontaneous. Bulk $Ni_5P_4(s)$, however, is only stable with respect to dissolution for -0.16 V $\geq U \geq$ -0.48 V at pH $= 0$, whereas $Ni_2P(s)$ is stable for -0.11 V $\geq U \geq$ -0.66 V. This prediction agrees with an experiment where it has been shown that applying negative potential ($\approx$ -0.2 V *vs.* RHE) suppresses degradation of Ni_5P_4. (245) For positive potentials, *i.e.* $U \approx$ 0.3 V *vs.* RHE, the compound dissolves with a rate of 1 ng/s/cm^2. (245) Therefore, the performance of Ni_5P_4 is expected to degrade over time for potentials higher than the stabilizing potential. Therefore, the lowest HER overpotential for $Ni_5P_4(000\bar{1})$ is -0.16 V, as the reaction is limited by the stability of bulk $Ni_5P_4(s)$. From this, we

see that the $(000\bar{1})$ surfaces of Ni_5P_4 are more HER active than that of $Ni_2P(0001)$, in agreement with experimental reports. (34) If one were only to consider thermodynamic barriers associated with the catalytic cycle, $Ni_5P_4(000\bar{1})$ would still have a lower overpotential (-0.07 V) than $Ni_2P(0001)$ (-0.14 V). Since we find that the adsorption energy of H on Ni_5P_4 is nearly thermoneutral, *i.e.* -0.06 eV, the intrinsic activity of Ni_5P_4 should be comparable to that of Pt. However, while it is experimentally found that the overall activity of nanocrystalline Ni_5P_4 approaches that of Pt on an electrode-basis, the surface-area-normalized turnover frequency (TOF) of Pt, *i.e.* performance on the basis of intrinsic activity (irrespective of morphology), is two orders of magnitude higher than that of Ni_5P_4. (34) This apparent inconsistency between the calculated catalytic overpotential and the experimentally measured TOF can be explained by the limited aqueous stability of bulk Ni_5P_4, requiring an applied potential $U \leq$ -0.16 V (higher potentials are expected to degrade the phosphides' performance). This reconciles the "apparent" intrinsic activity of Ni_5P_4 being below that of Pt, although we predict that purely on the basis of catalytic activity of the active surface, $Ni_5P_4(0001)$ is comparable to that of Pt. Synthetic methods, *e.g.* chemical doping, that would allow for Ni_5P_4 to be stable near 0 V *vs.* SHE are therefore expected to lead to higher performance for the phosphide.

Since our calculations indicate that the P_3-hollow is the most active site for HER catalysis, we propose high-throughput searching for materials that express this motif. P_3-hollows form on $Ni_5P_4(s)/Ni_4P_3(000\bar{1})$ because of the high P content in Ni_5P_4 relative to other bulk nickel phosphides and the ability of bulk $P(s)$ to form stable clusters. (67) Clustering behavior in nonmetals is not unique to P, as $S(s)$ also exhibits many different allotropes with various clustering geometries. As such, we believe that multi-nonmetal sites in general may hold promise for HER catalysis and should

be a focal point for researchers studying the HER. Another way to improve the efficiency of nickel phosphides for HER therefore would be to subtly modulate the metal-P bond strength so as to indirectly shift the free energy of H adsorption toward thermoneutrality. This opens a clear path for materials design of new materials to improve HER activity by tuning metal-P bond strength via lattice strain (*e.g.* epitaxial film growth) or low concentration doping/ion exchange with other elements. To date, however, there has only been one study of the relationship between Fe and Co-doping on the H adsorption free energy. (27) Since we find P to be the active site for HER on Ni_2P and Ni_5P_4, we recommend anionic substitutions for P as a more straightforward path to tune the catalytic activity, (246–248) in contrast with the transition metal substitution approach proposed recently. (27) For example, the presence of the more electronegative S can directly (through the formation of S-H) or indirectly (through modulation of the Ni-P bond strength) affect the binding of H onto the surface. Given that Ni_2P is a hydrodesulfurization catalyst (184), surface substitution with S should be possible. Nickel phosphide HER catalysts are reported to have nearly quantitative Faradaic efficiencies. (32; 249–251) Therefore, the formation of reduction side products from the electrolyte leading to the formation of reduced sulfur species is assumed not to occur and not included in our models. Further studies are necessary to develop a robust theory of chemical bonding between metal, P, and H in transition-metal phosphides, which could ultimately provide systematic guidance for designing improved HER catalysts.

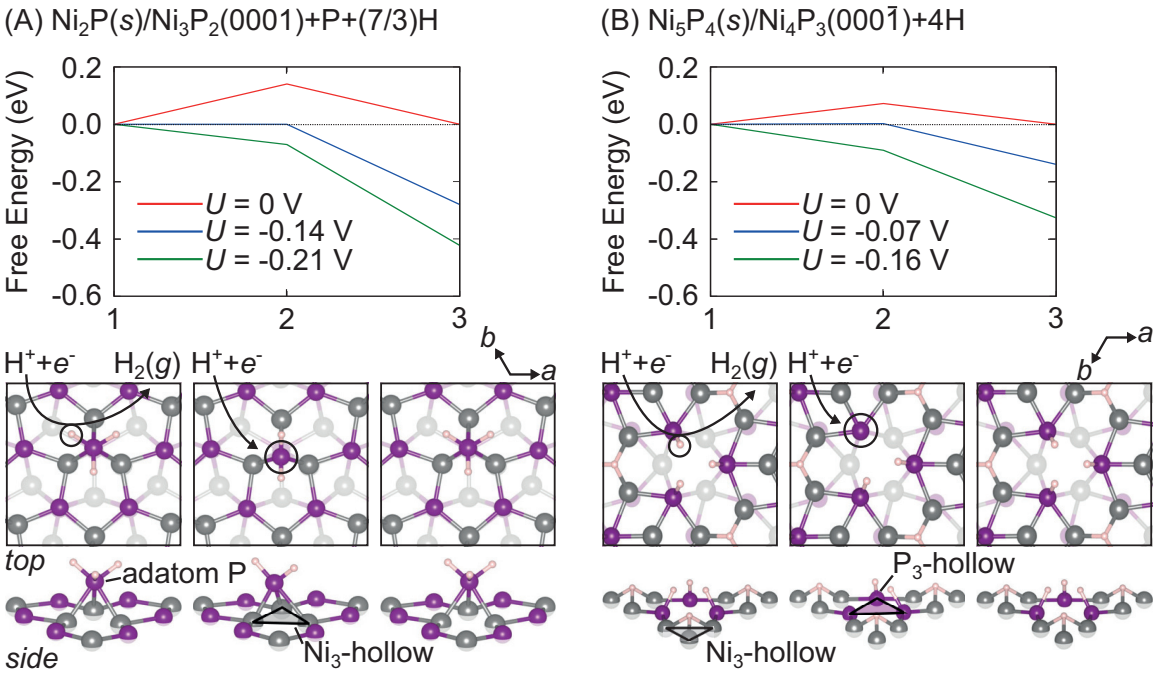

Figure 4.4: Free energy and structures of intermediates in the HER for (A) $Ni_2P(0001)$ and (B) $Ni_5P_4(000\bar{1})$ in equilibrium with 1 M Ni^{2+} or Ni(s), and 1 M PH_3 or 1 M H_3PO_4 at 298.15 K and pH = 0. The blue line corresponds to minimum overpotential to make the reaction spontaneous. The green line, however, corresponds to minimum overpotential to make the reaction spontaneous and ensure catalyst stability.

4.4. Conclusions

The $(000\bar{1})$ surface of Ni_5P_4 provides lower HER overpotentials and therefore greater HER activity than $Ni_2P(0001)$, which can be attributed to the thermodynamics and structure of surface P. For $Ni_2P(0001)$, the most stable aqueous surface reconstructions are $Ni_2P(s)/Ni_3P_2(0001)+n\mathrm{H}$ ($n = 1, 5/3$) and $Ni_2P(s)/Ni_3P_2(0001)+\mathrm{P}+(7/3)\mathrm{H}$, with the latter having an HER overpotential of -0.21 V. For Ni_5P_4, the $(000\bar{1})$ facet is more catalytically active than (0001), and the most stable aqueous reconstruction is $Ni_5P_4(s)/Ni_4P_3(000\bar{1})+n\mathrm{H}$ ($n = 4, 14/3$). This surface has the lowest overpotential for HER at -0.16 V. P, and not Ni, is the most active site, with adatom P and P_3-hollows providing low overpotential HER via the Volmer-Heyrovsky mechanism on Ni_2P and Ni_5P_4 (0001) surfaces. The P_3-hollow site on $Ni_5P_4(s)/Ni_4P_3(000\bar{1})$, which is present at low overpotential and offers nearly optimal H adsorption, is the origin of the superior catalytic activity of Ni_5P_4. The structural flexibility of P, *i.e.* its ability to form surface adlayers (adatom P) and in-plane multi-P clusters (P_3-hollow), provides a new frontier for improving the catalytic activity of transition-metal phosphides by embarking on high-throughput computational searches for catalysts that express these motifs and modulating the interaction strength between the metal and phosphorus via strain and surface doping.

CHAPTER 5 : Tuning the H_2 evolving activity of Ni_2P via surface nonmetal doping-generated chemical pressure: a joint first principles and machine learning study

5.1. Introduction

The discovery of highly active, noble-metal-free catalysts for the hydrogen evolution reaction (HER) is crucial for the development of economical water-splitting fuel storage technologies. There have been many candidates proposed in the last two decades, most notably MoS_2 (193; 202) and Ni_2P (32; 33). Recently, we found that the bulk-like Ni_3P_2 termination of the $Ni_2P(0001)$ surface, which is stable at modest electrochemical conditions, *i.e.* reducing potentials greater than -0.21 V *vs.* the standard hydrogen electrode (SHE) and $p\mathrm{H} = 0$, is not catalytically-active because the Ni_3-hollow site binds H too strongly ($\Delta G_{\mathrm{H}} = -0.45$ eV). (38) Upon the application of -0.21 V *vs.* SHE, however, the surface becomes enriched with P adatoms, which provide nearly thermoneutral H adsorption and consequently catalytic activity toward the HER. (38)

In order to overcome the inactivity of the Ni_3-hollow site, attempts have been made to tune the HER activity of Ni_2P by doping with different transition metals such as Co (252–254), Fe (255), Mn (256), and Mo (257). There have been no attempts, however, to dope Ni_2P with nonmetals, despite the host of stable binary Ni-nonmetal compounds that exist in nature and are catalytically active toward water splitting such as Ni_3N (258), Ni_3Se_2 (259), and Ni_3C (260). Furthermore, there have been no studies, experimental or theoretical, that investigate the effect of surface doping on

the HER activity of Ni_2P, as most studies have considered bulk doping. Here, we study the influence of surface nonmetal doping on the surface structure, charge states, and HER activity of $Ni_2P(0001)$ using density functional theory (DFT) calculations. We find that the Ni-Ni bond length is a robust descriptor for the HER activity of $Ni_2P(0001)$ using machine learning based on regularized random forests (159). This chapter outlines a transferable approach for the use of machine learning to extract descriptors from DFT structural and charge data.

5.2. Methods

5.2.1. First Principles Calculations

Spin-polarized DFT calculations were performed using the Quantum Espresso code (version 5.1) (240). Optimized, norm-conserving, designed nonlocal pseudopotentials were used to replace the nuclear Coulomb potential plus core electrons with a smoother, effective potential (105; 106). The valence electron wavefunctions were expanded in a plane-wave basis with an energy cutoff of 50 Ry. The generalized gradient approximation (GGA) of Perdew, Burke, and Ernzerhof (PBE) (87) was used to calculate the exchange-correlation energy. Grimme's semiempirical DFT-D2 method (129; 261) was used to include dispersion interactions, which are generally important for accurately modeling catalytic processes. (241; 242) We choose the DFT-D2 method because it shows excellent agreement with higher-level electron correlation methods, *i.e.* Møller-Plesset perturbation theory (MP2) and coupled-cluster singles and doubles (CCSD), for H adsorption energies. (262; 263)

We modeled the HER on an eight-layer, periodic slab of Ni_2P. The dimensions of the slab were $a = b = 10.07$ Å and $c = 39.57$ Å, and $\alpha = \beta = 90°$ and $\gamma = 120.08°$. The width of the vacuum region was 26.24 Å. A $(\sqrt{3} \times \sqrt{3})R30°$ surface supercell was

used so that we could model fractional surface concentrations of nonmetal dopants. A Γ-centered, $3 \times 3 \times 1$ grid of k-points was used to sample the Brillouin zone. During geometry relaxations, the bottom four layers of the slab were fixed in their bulk configurations.

5.2.2. Machine Learning

Regularized random forests (RRFs) were trained using the caret package (version 6.0.77) for R (version 3.2.5). (153) Processed data and R scripts for the machine learning can be found in Appendix C. Three-fold cross-validation (CV) was performed to improve our prediction of the "out-of-sample" error. At each step of the training process, 10 descriptors were randomly selected and a regularization value of 0.1 and an importance coefficient of 0.75 were applied. More details on the DFT calculations and machine learning can be found in Appendix C.

5.3. Results and Discussion

5.3.1. Surface Structure and Doping Scheme

Bulk $Ni_2P(s)$ has two alternating layers along the [0001] axis with compositions of Ni_3P and Ni_3P_2. Fig. 5.1A shows the structure of the $Ni_3P_2(0001)$ termination. The surface has one Ni_3-hollow site per unit cell (three per $\sqrt{3} \times \sqrt{3}R30°$ supercell). In an aqueous electrochemical environment at $U = 0$ V *vs.* SHE and pH $= 0$, each Ni_3-hollow site binds H strongly ($\Delta G_H = -0.45$ eV, $n_X = 0$ in Fig. 5.1B). Each Ni bonds with two P in the surface layer and one P in the subsurface layer. This makes for a total of six symmetry-equivalent P atoms surrounding the Ni_3-hollow site in the surface layer, which are numbered from 1 through 6 in Fig. 5.1A. We replace these surface P sites with nine different nonmetals (As, B, C, N, O, S, Se, Si, and Te) and

varied the number of dopants (n_X) from 1 to 6. We found that the S dopants prefer being separated at $n_S = 2$, *i.e.* S at positions 1 and 2 is more favorable than when they are sharing a common Ni (at positions 1 and 5) by 0.34 eV. Thus, for other dopants, and at higher doping concentrations, the maximum separation between dopants is maintained. The indices on P in Fig. 5.1A correspond to the preferred sequence of substitution of the P atoms. For example, at $n_X = 3$, dopants are substituted at sites 1, 2, and 3. While the Ni and P sites are symmetrically-equivalent initially, doping breaks this symmetry.

5.3.2. Effect of Doping Concentration on H Adsorption and Dopant Substitution

H Adsorption

The effect of surface dopant identity and concentration on ΔG_H at the Ni_3-hollow site is shown in Fig. 5.1B. We observe three distinct trends corresponding to (1) As and Si, (2) 2*p* nonmetals (B, C, N, and O), and (3) chalcogenides (S, Se, and Te). The first set does not differ appreciably from the undoped surface (dashed, light blue line marked "undoped"). ΔG_H is relatively constant with respect to changes in the surface dopant concentration. Conversely, 2*p* nonmetals (set 2) have a dramatic effect on ΔG_H, which generally increases with increasing n_X. Nearly thermoneutral H adsorption is possible for $n_X = 2-3$ (50% substitution) and, above that, H adsorption is no longer favored. The chalcogenides (set 3), on the other hand, have an intermediate effect on ΔG_H. From $n_X = 1$ to 3, ΔG_H increases at about the same rate for each chalcogenide. From $n_X = 4$ to 6, however, ΔG_H decreases, with the attenuation being more pronounced for S and Se than Te. The maximum ΔG_H achieved is -0.11 eV for S at $n_X = 3$. Therefore, doping the $Ni_3P_2(0001)$ surface of Ni_2P with 2*p* nonmetals (B, C, and O) or chalcogens at $\approx 50\%$ ($n_X \approx 3$) can substantially improve the HER activity of the

Ni_3-hollow site.

Dopant Substitution

Next, we evaluate the stability of different doping configurations. To do this, we calculate the free energy of substitution (ΔG_{sub}) relative to the computationally favorable phases under reducing conditions $U = 0$ V *vs.* SHE and pH $= 0$ of P and each dopant (H_3PO_4, As, SiO_2, H_3BO_3, CH_4, NH_4^+, H_2O, H_2S, H_2Se, and Te) (264) for $n_{\text{X}} = 1$ to 6. Details and an example calculation of ΔG_{sub} for can be found in Appendix C (see Table C.2). Fig. 5.1C shows that ΔG_{sub} does not depend strongly on n_{X}; however, a slight increase is observed for As, the chalcogens, C, N, and O. Substitution is exergonic for only five of the nine nonmetals: As, O, and the chalcogens. The others, period two nonmetals and Si, are significantly endergonic. The only set of nonmetals that both enhances the HER activity of the Ni_3-hollow site and stabilizes the surfaces is the chalcogens. O is an exception in set 2 as it substitutes spontaneously with P at $n_{\text{X}} = 2$ and 3. This has a negative effect on catalysis because at $n_{\text{X}} = 3$, where O substitution is most favorable, $\Delta G_{\text{H}} = 0.22$ eV, which is too weak for facile hydrogen evolution. However, since ΔG_{sub} for the heavier chalcogens are much more favorable, substitution with S, Se, and Te will be able to inhibit the introduction of O in the surface. Such substitutions can thus suppress catalytically degenerative oxygenation of the surface, highlighting another advantage of nonmetal doping. As such, we propose surface doping with S, Se, and Te as a very promising route for improving the HER activity of $Ni_3P_2(0001)$.

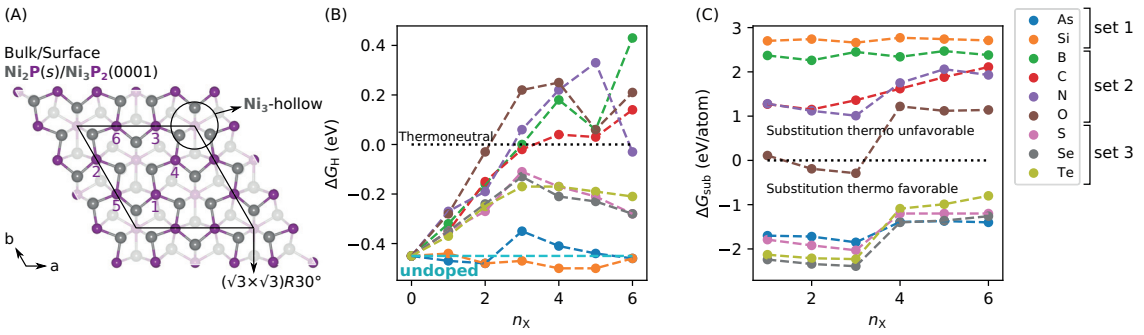

Figure 5.1: (A) Structure of $Ni_3P_2(0001)$ surface of Ni_2P showing the $(\sqrt{3}\times\sqrt{3})R30°$ supercell. The Ni_3-hollow sites, which bind H, are shown. The indices on P atoms indicate the preferred sequence of substitution with dopants. Free energy of (B) H adsorption and (C) dopant substitution as a function of the surface dopant concentration. $\Delta G_H = 0$ is referred to as "thermoneutral" H adsorption. ΔG_H for the undoped surface is labeled and denoted by a dashed, light blue line. The spontaneity of dopant substitution is labeled and indicated by a dotted black line.

5.3.3. Exploratory Data Analysis

Charge Descriptors

In order to better understand the trends in ΔG_{H} and ΔG_{sub} with respect to n_{X}, we first perform a rudimentary exploratory analysis of charge descriptors. Fig. 5.2A shows nearly linear trends between the average Ni residual charge ($\langle q_{\mathrm{Ni}} \rangle$) and number of dopants (n_{X}). Their slopes can be interpreted as the direction of electron transfer between Ni and the dopant. For example, positive slopes correspond to electron transfer from Ni to the dopant, thereby oxidizing Ni (the electron donor) and reducing the nonmetal (the electron acceptor). The opposite is true for negative slopes. $\langle q_{\mathrm{Ni}} \rangle$ does not correlate strongly with ΔG_{H}. For example, S and Se have very similar trends in ΔG_{H} (see Fig. 5.1B). However, S and Se cause opposite shifts in $\langle q_{\mathrm{Ni}} \rangle$. Additionally, there are no significant changes in their $\langle q_{\mathrm{Ni}} \rangle$ trends that coincide with the maximum in ΔG_{H} at $n_{\mathrm{X}} = 3$. Therefore, $\langle q_{\mathrm{Ni}} \rangle$ is a poor descriptor of ΔG_{H}.

Important Descriptors from Machine Learning

In order to rationalize the trends in ΔG_{H}, we search for other structural and charge descriptors. For each DFT-relaxed structure, we compile Ni-Ni bond lengths, Ni-Ni-Ni bond angles, Löwdin charges, elemental data (mass number, atomic weight, and atomic radius), summary statistics (mean and standard deviation), and other geometric parameters (perimeter and area of Ni_3-hollow sites). Note that the descriptors we chose involve only the H adsorption site, *i.e.* the Ni_3-hollow site, and the dopants. Descriptors involving surface P atoms are deemed unnecessary because they don't directly participate in bonding, and changes in the Ni-P bond lengths are in fact already included in the contraction or expansion of the Ni_3-hollow sites. Removing

such potentially redundant descriptors keeps the data simple. Our data are obtained from surfaces without H, because we want to be able to predict HER activity based on intrinsic surface properties. In total, the data set has 55 observations (structures) and 30 variables (29 descriptors and ΔG_{H}). Machine learning is becoming increasingly popular for semi-automated and quantitative discovery of data correlations in chemistry and materials science. (265–267) Using these data, we trained an RRF, yielding a root-mean-squared error (RMSE) of 0.09 eV. Fig. 5.3A shows that the data very closely straddle the perfect correlation line. Note that we performed three-fold CV instead of randomly splitting the data into one training and one test set (see Fig. C.1 in Appendix C). In general, k-fold CV splits the data into k sets and averages the models generated by training on $k-1$ of the sets and testing on the other. This method gives better estimates of the "out-of-sample" error, which in our case refers to the RMSE in the prediction of ΔG_{H}, than randomly train/test splitting.

The architecture of RRFs allows for the calculation of feature importances. These measure the relative importance of different descriptors in describing the HER activity of nonmetal-doped $Ni_2P(0001)$ surfaces. Importance is defined as the normalized ability of a descriptor to separate the data based on ΔG_{H}. Fig. 5.3B shows the top ten, most important descriptors from the 29 included in our data set. The top two descriptors are a particular Ni-Ni bond length (whose constituent atoms are distinguished by their distance from the first doping site, see Fig. 5.3C) and the average Ni-Ni bond length ($\langle$Ni-Ni$\rangle$). Of the top ten descriptors, seven are related to the geometry of the Ni_3-hollow site. Other important features include the charges of dopants at $n_{\mathrm{X}} = 2$ and 3 (q_{X2} and q_{X3}), and the standard deviation of the dopant charges (σ_{qX}). While in principle, it would be straightforward to include more complex descriptors like the electronic DOS and its moments (*e.g.* the *d*-band center (268; 269)

of Ni), this is shown to be unnecessary, however, because the simple and intuitive Ni-Ni bond length proves to be quite descriptive of HER activity. Further, the fact that the atomic charges exhibit poor correlation indicates that such metrics that depend on electronic partitioning will also likely be unimportant.

Structural Descriptors

Having identified ⟨Ni-Ni⟩ as a good descriptor for ΔG_{H}, we more closely examine their correlation. Fig. 5.2B shows the effect of surface doping on ⟨Ni-Ni⟩ as a function of n_{X}. Like ΔG_{H}, ⟨Ni-Ni⟩ is relatively unaffected by doping with As and Si. The period two nonmetals, however, induce a significant expansion in the ⟨Ni-Ni⟩. This correlates quite strongly (Pearson's $r \geq 0.41$, $r = -1$ or 1 for perfect negative and positive correlation, respectively) with ΔG_{H}, which also increases dramatically with increasing n_{X}. The chalcogens (set 3) show two regimes of change in ⟨Ni-Ni⟩, much like their trends for ΔG_{H} and ΔG_{sub}. For set 3, at lower surface doping concentrations (*i.e.* $n_{\mathrm{X}} = 1$ to 3), the ⟨Ni-Ni⟩ increases with the increase for S being largest and Te the smallest. At higher surface doping concentrations (*i.e.* $n_{\mathrm{X}} = 4$ to 6), ⟨Ni-Ni⟩ plateaus. While there is an apparent moderate dependence of ΔG_{H} or ⟨Ni-Ni⟩ on n_{X} seen in Figs. 5.1B and 5.2B, it only became evident that Ni-Ni bond lengths are good descriptors for HER activity through the training of the RRF model. Since the influence of surface nonmetal doping on both the geometry of the Ni_3-hollow site and ΔG_{H} are similar, this implies a chemical pressure-like effect (270) that can be summarized as follows. Nonmetal surface doping effectively acts like mechanical pressure, expanding or compressing the Ni_3-hollow site.

This effect can be rationalized as follows. As the Ni_3-hollow site expands, the Ni-H bonds that form will have to stretch if H is to remain at the center. Upon sufficient

expansion, however, H must reduce/sever its interaction with one or two Ni atoms to form an optimal Ni-H bond length, thereby weakening its adsorption strength (see Fig. C.3 in Appendix C for an example of this). The HER can be broken down into two steps, H adsorption and H_2 desorption. The former is called the Volmer reaction, and its rate is proportional to the strength of H adsorption. H_2 desorption can follow either the Tafel or Heyrovsky mechanism. The rates for both of these reactions are inversely proportional to the strength of H adsorption. In order to maximize the rate of the HER, a compromise, often called the Sabatier principle (271), must be made between H adsorption and H_2 desorption. This compromise is struck at $\Delta G_{\mathrm{H}} = 0$, which is referred to as thermoneutral H adsorption. As a final comment, the Brønsted-Evans-Polanyi (BEP) principle states that there is a linear relationship between the kinetic barrier and the free energy of a reaction. Therefore, by weakening H adsorption at the Ni_3-hollow site via dopant-induced tensile strain, the kinetic barrier for HER is also decreased and its rate is increased.

Local charges play a minimal role because the surface is metallic (67) and therefore can easily provide the requisite free charge to stabilize adsorption. The effect of doping on the surface strain is nonlinear with the dopant atomic radius because the induced strain is a complex function of the relative electronegativity of the constituent atoms, valence, concentrations of the dopants, and coordination chemistry. Naturally, dopants that form shorter Ni-X bonds will expand the Ni_3-hollow site, while those that form longer Ni-X bonds will cause the Ni_3-hollow site to contract. This explains the more drastic effect of $2p$ nonmetals on the adsorption of H due to the Ni_3-hollow site's expansion. However, the mechanical effect of X on Ni-X bonding is also dependent on the dopant concentration, hence the observed nonlinear dependence of $\langle \mathrm{Ni} - \mathrm{Ni} \rangle$ on n_{X} (Fig. 5.2B). These explain why the dopant atomic radius appears to be a less

important descriptor for ΔG_{H}.

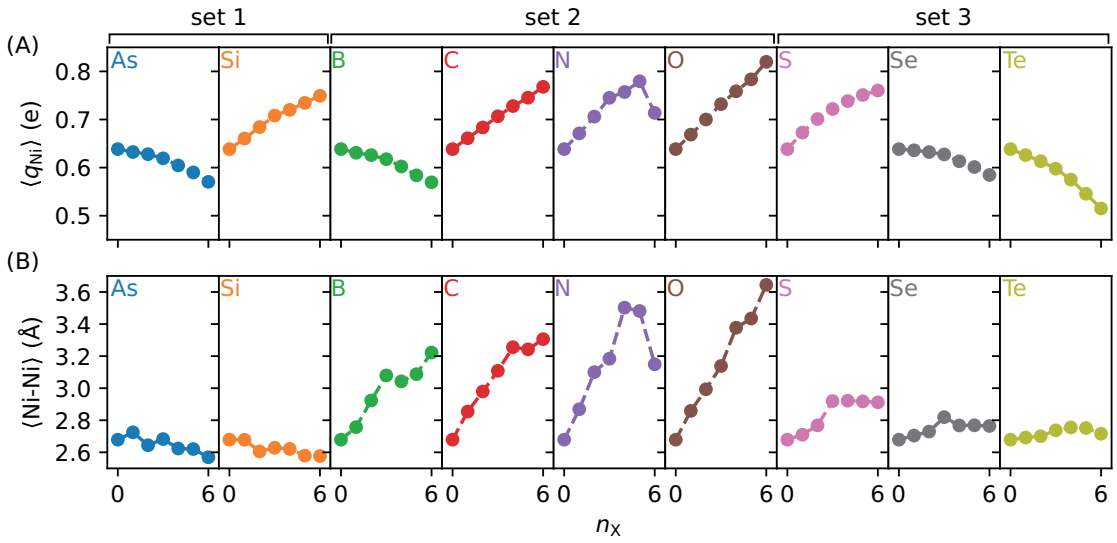

Figure 5.2: Effect of dopant and surface concentration on the (A) average Ni residual charge ($\langle q_{\mathrm{Ni}} \rangle$) and (B) average Ni-Ni bond length ($\langle$Ni-Ni$\rangle$).

Chemical Pressure Proof of Concept

In order to confirm this machine learning insight, we apply mechanical pressure to the Ni_3-hollow site. The ⟨Ni-Ni⟩ was compressed and expanded by fixing the internal coordinates of the surface Ni atoms and allowing the other surface atoms to relax during DFT geometry optimization. Note that the lattice constants were fixed. Fig. 5.3D shows that applied mechanical pressure (orange points and dashed line) induces the same change in ΔG_{H} as does chemical pressure via nonmetal doping (blue points). Therefore, it is not the electronic character of the nonmetal dopants but rather the structural distortion they induce on the surface that modulates the HER activity of the $Ni_3P_2(0001)$ surface. Note that our calculations indicate that ⟨Ni-Ni⟩ contracts by ≈ 0.1 Å upon H adsorption. This means that since the internal coordinates of Ni are fixed even upon H adsorption, unlike in the doping case, we are likely overestimating ΔG_{H}, and thus the orange line represents an upper limit for the mechanical pressure effect. If we manually adjust ⟨Ni-Ni⟩ by 0.1 Å (Fig. 5.3D green dotted line), the agreement between mechanical and chemical pressure improves.

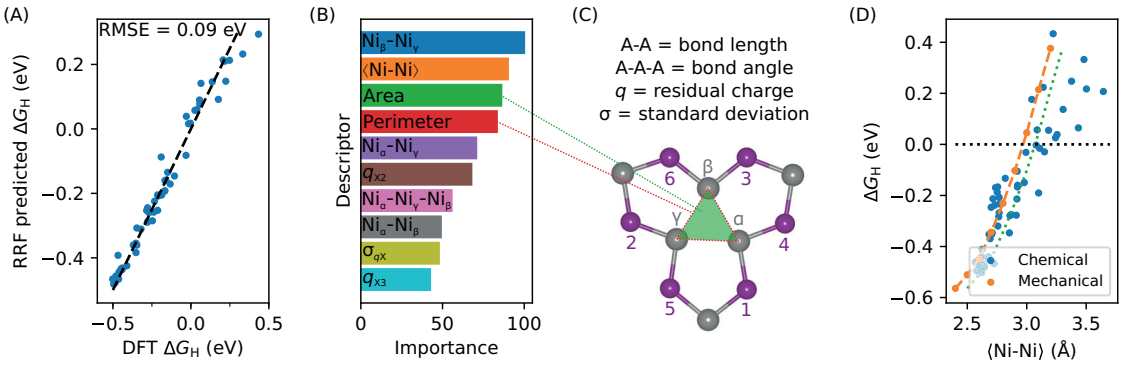

Figure 5.3: (A) ΔG_{H} predicted by RRFs *vs.* DFT. Black-dashed line corresponds to perfect agreement. (B) Relative importance of descriptors calculated from RRF model. Only the top 10 features are shown (see Fig. C.2 in Appendix C for full list). (C) Definition of descriptors in (B). We label the three Ni atoms α, β, and γ based on their distance from the first doping site. (D) Effect of average Ni-Ni bond length on ΔG_{H} as induced by chemical pressure and mechanical pressure. Chemical pressure was induced by surface nonmetal doping and mechanical pressure by fixing the positions of surface Ni atoms. Green dotted line adjusts for Ni-Ni bond contraction upon H adsorption in the mechanical case. According to the mechanical pressure calculations, the ideal Ni-Ni bond length for HER is between 2.97 and 3.07 Å.

Perspectives

Based on the mechanical pressure calculations, $\langle$Ni-Ni$\rangle \approx 2.97 - 3.07$ Å should produce thermoneutral H adsorption and thus the optimal intrinsic activity for an HER electrocatalyst expressing the Ni_3 motif. The optimal bond length should be used in high-throughput searches to screen for bulk binary Ni-nonmetal compounds. An alternative would be to study mixed nonmetal doping, *e.g.* doping with both S and Se. We anticipate that undoped and doped transition metal phosphides with bulk crystal structures similar to $Ni_2P(s)$, such as $Fe_2P(s)$, $Co_2P(s)$, (Fe,Co)P(s), and (Ni,Co)P(s), will also exhibit chemical pressure-driven enhancement of HER to varying degrees. Our study demonstrates that a comprehensive investigation of surface nonmetal doping for a variety of single and mixed transition metal phosphides will reveal promising candidate materials with nearly ideal intrinsic activity toward the HER.

In these kinds of doping studies, we should not lose sight of the thermodynamics and kinetics of dopant incorporation and segregation within the bulk, as this will provide valuable information regarding the feasibility of synthesizing catalysts. We have shown that certain dopants may stabilize the surface with respect to dissolution, *e.g.* doping with chalcogens or As.

Our current data set for machine learning is specifically obtained for the Ni_2P surface that expresses the Ni_3 motif, and thus the method will work best in predicting perturbations within the structural framework of Ni_2P. A great example of extending this model would be the examination of doping with transition metals, as mentioned above, in cases where this only causes minimal changes to the underlying atomic structure of Ni_2P. A model that is transferable across different bulk transition-metal

phosphides would of course require additional structure-specific inputs during training. Although the RRF is trained using only a subset of dopant arrangements, having verified the predictive power of Ni_3-structure-based descriptors, the RRF model should also be able to predict the ΔG_H of other dopant configurations with high precision and accuracy.

Unlike $\langle$Ni-Ni$\rangle$, $\langle q_{Ni}\rangle$ does not correlate strongly with ΔG_H. Even though the nonmetal dopants substitute at spectator sites, it is contrary to common understanding of dopant effects that their electronic structure does not play a large role in determining HER activity. This could be explained by the fact that $Ni_2P(s)$ is metallic and therefore charge partitioning between the Ni and nonmetal components is less well defined. This hypothesis is corroborated by the results in Fig. 5.2A, which show that the average charge on Ni only changes by a small fraction of the charge of an electron from low to high surface doping concentrations. The structural parameters of the Ni_3-hollow site are more sensitive to changes in the surface electronic structure and therefore they are able to more accurately capture the trends between n_X and ΔG_H. This lends further support to our claim that chemical pressure is the key driving force behind the enhanced HER activity of doped $Ni_3P_2(0001)$. The connection between strain and catalytic activity has been explored in the literature, especially with regard to substrate-induced (272; 273) and electrochemically-induced (274) strain. For example, compressive strain was demonstrated to enhance the oxygen reduction reaction on dealloyed Pt-Cu and Pt nanoparticles via enhancement of the binding of intermediate oxygenated adsorbates, (273; 274) while tensile strain was shown to stabilize CO and O chemisorption and CO dissociation on Ru(0001). (272) Here, expansion of the Ni_3 site (local, chemically-induced tensile strain) reduces the affinity of H due to the reduced H coordination. The most promising catalytic system that

we discovered was chalcogen-doped $Ni_3P_2(0001)$. For S ($n_S = 3$), the HER overpotential is -0.11 V *vs.* SHE. This is much lower than that of the Ni_3-hollow site in the absence of dopants (-0.45 V *vs.* SHE). (38) This overpotential is similar to that of the P-enriched, nonstoichiometric reconstruction (-0.07 V *vs.* SHE) but this surface is only accessible upon applying -0.21 V *vs.* SHE. (38)

5.4. Conclusions

In summary, we have demonstrated that surface nonmetal doping can significantly improve the HER activity of Ni_2P. We find that the Ni-Ni bond length is an effective descriptor for the HER activity of $Ni_3P_2(0001)$ of Ni_2P and hence can be used in a computationally efficient, high-throughput search to screen for promising Ni-nonmetal catalytic materials. We have shown how machine learning methodologies can be implemented in the catalyst design pipeline to automatically discover and rank the importance of structural and charge-based descriptors for HER. Machine learning is highly customizable in that many different model types can be selected (here we choose RRFs) and the number and types of descriptors is limited only by scientific creativity. We validated the results from our machine learning by applying mechanical pressure to compress and expand the Ni_3-hollow sites, which showed that the effects of chemical pressure via nonmetal doping and mechanical pressure are in excellent agreement. Our results strongly indicate that it is the induced local geometry of the Ni_3-hollow site and not the electronic character of the dopants that improves the HER activity of $Ni_3P_2(0001)$. We believe that this insight should spur both experimental and theoretical research in surface nonmetal doping of transition metal phosphides in the wider effort find ideal HER catalysts.

CHAPTER 6 : Surface crystal structure prediction using *ab initio* grand canonical Monte Carlo simulations

6.1. Introduction

Theoretical modeling of surface chemical and physical properties often involves making assumptions about the surface structure. However, the physical and chemical properties depend sensitively on these assumptions. The simplest starting point for constructing a surface model is to select a particular facet and then to identify bulk-like terminations from the layering pattern normal to that surface. This approach, however, does not take into account the fact that many bulk-terminated surfaces undergo reconstruction in order to chemically passivate surface bound charges and/or saturate surface atom coordination. (45; 49; 190; 191; 275–279) Therefore, the ideal approach involves an exhaustive exploration of all possible surfaces and their reconstructions.

Such an undertaking has two main drawbacks: its computational cost can be prohibitive, and the phase space of surface structures is vast and sometimes surprising. Recently, progress has been made toward overcoming these drawbacks by using machine learning to more efficiently traverse surface phase space. For example, genetic algorithms have been developed that programmatically mate different surfaces to explore lower-symmetry phases. (62; 280; 281) Additionally, Gaussian process regression has been employed to learn intermediate surfaces, *i.e.* those that are a mixture of phases from the training set, thereby reducing the number of first principles calculations necessary. (66) Despite the power of these methods, their main goal is to minimize the surface energy, and they accomplish this using effective but poten-

tially unphysical structural transformations, thus rendering them unable to provide mechanistic information about the natural evolution of the surface.

A simpler and more physically motivated way to explore surface phase space is grand canonical Monte Carlo (GCMC). In GCMC simulations, a system is in contact with both thermal and chemical potential reservoirs, thus allowing fluctuations in the temperature and number of particles. Historically, this technique has been used to study adsorption isotherms: molecules on metals, (282) metal-organic frameworks, (283–285) carbon-based materials, (286; 287) zeolites, (288; 289) ionic liquid, (290) and activated carbon. (291) GCMC has also been applied to study the bulk phase diagrams of liquids, (292) their mixtures, (293) alloys, (294–296) and fluids, (297) and solvation phenomena. (298; 299) In principle, GCMC can be used to generate a collection of surface structures consistent with a predefined temperature and set of chemical potentials of the constituent elements. An application of GCMC to the prediction of surface reconstruction, despite its simplicity and elegance, has never been attempted.

In order to evaluate the efficacy of GCMC in predicting surface phase diagrams, tests on a well-understood yet complicated material must be performed. One such material that fits these criteria is Ag, which plays an important role in plasmonics, (300; 301) catalysis, (302; 303) and medicine. (304; 305) Since the 1970s, many versions of the Ag(111) surface have been proposed, supported, rejected, and accepted. (55; 56; 277–279; 306–309) Early on, low-energy electron diffraction (LEED) and X-ray photoelectron spectroscopy (XPS) measurements suggested that an $Ag_2O(111)$ overlayer with $p(4 \times 4)$ surface periodicity grows on Ag(111) due to their nearly matching lattice constants. (306; 307) With the advent of scanning tunneling microscopy (STM) and the reemergence of *ab initio* thermodynamics, a host of new structures

were proposed, including Ag-deficient and O-enriched variants of the Ag_2O overlayer, (55; 277; 308) an $Ag_{1.2}O$ cloverleaf-like overlayer, (277) and, most recently, an overlayer consisting of Ag_9 islands each connected by two O atoms. (56; 278; 309) Additionally, surface structures with many other periodicities have been observed experimentally, such as a $c(4 \times 8)$ overlayer, which possesses stripes of base-connected Ag_3O_4 triangular pyramids; to date, this $c(4 \times 8)$ pattern offers the lowest surface free energy (for $\Delta\mu_{Ag} = 0$ eV and -0.64 eV $\lessapprox \Delta\mu_O \lessapprox$ -0.19 eV) of any Ag(111) reconstruction, as calculated from density functional theory (DFT). (279)

Here, we report the design of an algorithm and the development of a computer program that implements GCMC in the DFT software package Quantum ESPRESSO. (310) Our implementation of GCMC is open-source, portable, and requires a small number of user inputs. (311) We show that *ab initio* GCMC, with a small set of simple configurational biases, can independently (re)discover the key features of the oxidized Ag(111) surface phase diagram, which puzzled surface scientists for five decades. We also show that by analyzing the *ab initio* GCMC results with a machine learning model, we can understand and explain the relationships between different structural features and the surface energy. We propose *ab initio* GCMC as a flexible, general-purpose tool that not only facilitates the discovery of surfaces that are likely to be obtained under different conditions but also generates a rich data set that, upon interrogation, reveals the driving forces behind the formation of different surface structures.

6.2. Methods

6.2.1. Theory

We work in the grand canonical ensemble, where the chemical potential μ, volume V, and temperature T of the system are fixed. The partition function of the grand canonical ensemble is

$$Q(\mu, V, T) = \sum_{N=0}^{\infty} \frac{e^{\frac{\mu N}{k_B T}} V^N}{\Lambda^{3N} N!} \int d\vec{s}^N e^{-\frac{U(\vec{s}^N)}{k_B T}} \tag{6.1}$$

where k_B is the Boltzmann constant, N is the number of atoms, Λ is the thermal de Broglie wavelength, given by $\Lambda = \frac{h}{\sqrt{2\pi m k_B T}}$, h is the Planck constant, m is the mass of the atom, U is the potential energy, and $\vec{s}^N$ are the fractional coordinates of the atoms. The probability density corresponding to a particular configuration $(\vec{s}^N; N)$ is

$$q_{\mu VT}\left(\vec{s}^N; N\right) \propto \frac{e^{\frac{\mu N}{k_B T}} V^N}{\Lambda^{3N} N!} e^{-\frac{U(\vec{s}^N)}{k_B T}} \tag{6.2}$$

There are three different types of actions in unbiased GCMC simulations: move the existing particles in the system, add particles to the system, and remove particles from the system. To ensure that the simulation satisfies detailed balance, the acceptance probability for each action must satisfy

$$q_{\mu VT}(1)\,\alpha(1,2)\,P(1,2) = q_{\mu VT}(2)\,\alpha(2,1)\,P(2,1) \tag{6.3}$$

where 1 and 2 represent configurations $\left(\vec{s}_1^{N_1}; N_1\right)$ and $\left(\vec{s}_2^{N_2}; N_2\right)$, respectively. $\alpha(1,2)$ is the probability of attempting a move from configuration 1 to 2 and $P(1,2)$ is the probability of accepting that move. Since $\alpha(1,2) = \alpha(2,1)$, the probability of

accepting an attempted "move" step (312) is

$$P_{\text{move}} = \min\left\{1, e^{-\frac{\Delta U}{k_B T}}\right\} \tag{6.4}$$

where ΔU is the change in potential energy. For an "exchange" step, if the probability of attempting an "add" and "remove" action are equal, *i.e.*

$$\alpha\left(N \text{ particles}, N+1 \text{ particles}\right) = \alpha\left(N+1 \text{ particles}, N \text{ particles}\right), \tag{6.5}$$

the acceptance rules are

$$P_{\text{add}} = \min\left\{1, \frac{V}{(N+1)\Lambda^3} e^{-\frac{\Delta U - \mu}{k_{\text{B}} T}}\right\} \tag{6.6}$$

and

$$P_{\text{remove}} = \min\left\{1, \frac{N\Lambda^3}{V} e^{-\frac{\Delta U + \mu}{k_{\text{B}} T}}\right\} \tag{6.7}$$

In order to focus on the growth of the surface in contact with thermal and chemical potential reservoirs, we replace the "move" action with a structural relaxation after "add" and "remove" actions. The bias introduced by structural relaxation can be countered by replacing the volume in the acceptance probability (see Eqs. 6.6 and 6.7) with an effective volume V_{eff}, as discussed in previous works. (313; 314) For the "add" action, we choose an element, each with an equal probability, and add it to the system. Instead of randomly selecting the position of the new atom, we include a configurational bias, which prevents the new atom from being too close ($r_{\min} = 1.5$ Å) or too far ($r_{\max} = 3$ Å) from the closest existing atom. If these criteria are not met, then we skip this step. This bias has little effect on the detailed balance because all of the configurations we rule out have very high energies and, practically speaking,

could never be accepted. For the "remove" action, we randomly choose an atom and remove it from the system. In order to further restrict the sampling to those phases relevant for surface growth, we added a constraint that atoms can only be inserted at or removed from positions near the top surface (see Fig. 6.1). A flowchart for our *ab initio* GCMC scheme can be found in Appendix D (see Fig. D.1) .

In this chapter, we study the Ag(111) surface and its reconstructions (see Fig. 6.1). In order to sample a variety of surface structures and compositions, we set the temperature of the simulations to 500 K and test a range of chemical potentials around the equilibrium between bulk Ag(s) and Ag_2O(s), for which $\mu_{\mathrm{Ag}} = \mu_{\mathrm{Ag}}^{\mathrm{eq.}} = G_{\mathrm{Ag}}$ and $\mu_{\mathrm{O}} = \mu_{\mathrm{O}}^{\mathrm{eq.}} = \frac{1}{2}\left(G_{\mathrm{Ag_2O}} - G_{\mathrm{Ag}}\right)$. The free energies of bulk Ag(s) and Ag_2O(s) can be approximately written as

$$\begin{aligned} G_{\mathrm{Ag}} &\approx U_{\mathrm{Ag}}^{\mathrm{DFT}} + \Delta G_{\mathrm{Ag}}\left(T\right) \\ G_{\mathrm{Ag_2O}} &\approx U_{\mathrm{Ag_2O}}^{\mathrm{DFT}} + \Delta G_{\mathrm{Ag_2O}}\left(T\right) \end{aligned} \tag{6.8}$$

where the temperature-dependent term is taken from experimental data. (315) We tested five different μ_O conditions such that $p_{\mathrm{O_2}}/p_{\mathrm{O_2}}^{\mathrm{eq.}} = \{10^{-10}, 10^{-6}, 10^{-2}, 1, 10^{2}\}$. Since the Ag/Ag_2O bulk phase boundary corresponds to relatively O-rich conditions, we choose three $p_{\mathrm{O_2}}$ lower and only one $p_{\mathrm{O_2}}$ higher than $p_{\mathrm{O_2}}^{\mathrm{eq.}}$. The change in the volume from V to V_{eff} in the acceptance probability can be interpreted as a change in the chemical potential, *i.e.* $\delta\mu = k_{\mathrm{B}}T \ln \frac{V}{V_{\mathrm{eff}}}$. $V/V_{\mathrm{eff}} \approx 10$ because the MC-inserted atoms can access only 10% of V, that which is not occupied by the existing atoms. $\delta\mu$ is approximately equal to a one order of magnitude change in partial pressure of O_2. Therefore, we can directly use V, instead of V_{eff}, since our simulation is performed over a range of chemical potentials, and it will not influence the result.

As is the convention in the literature, (316) we calculate the surface energy relative

to that of Ag(111),

$$\gamma^*_{\text{slab}} = \gamma_{\text{slab}} - \gamma_{\text{Ag(111)}} \tag{6.9}$$

where γ_{slab} is defined as

$$\gamma_{\text{slab}} = \frac{1}{A}\left(U^{\text{DFT}}_{\text{slab}} - n_{\text{Ag}}\mu_{\text{Ag}} - n_{\text{O}}\mu_{\text{O}}\right) \tag{6.10}$$

Here, A is the surface area and n is the number of atoms. A factor of two is missing from the denominator because the bottom layer of each slab is the same, *i.e.* Ag(111), and its contribution to γ^*_{slab} cancels out.

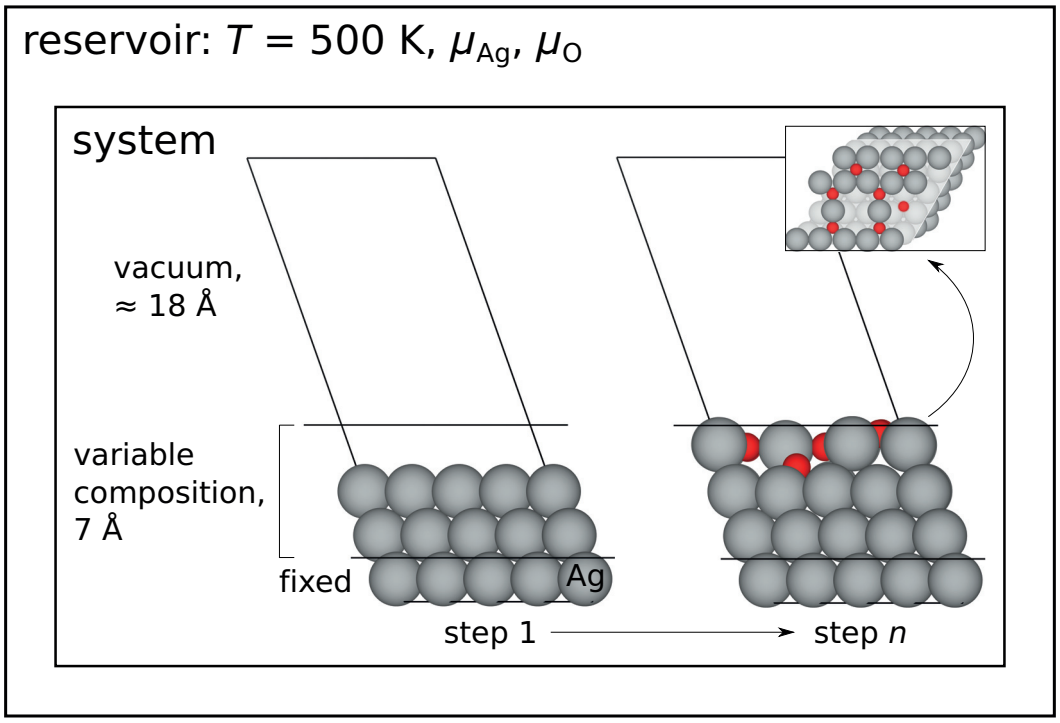

Figure 6.1: $p\,(4 \times 4)$ Ag(111) slab model for GCMC simulations. We set the temperature and the chemical potentials of Ag and O. The surface is three layers thick, with the bottom layer fixed and ≈ 18 Å of vacuum. Atoms are only added to or removed from the variable composition region, which extends from 3.5 Å below to 3.5 Å above the top layer of Ag.

6.2.2. Computational Details

DFT (72; 177) calculations were performed using Quantum ESPRESSO (version 6.2.1). (310) The generalized gradient approximation (GGA) of Perdew, Burke, and Ernzerhof (PBE) was used to treat electron exchange and correlation. (87) Designed nonlocal, (105) optimized, (106) norm-conserving pseudopotentials (103) were generated for Ag and O using OPIUM. (179) We used $5s$, $5p$, and $4d$ as the valence states for Ag and $2s$ and $2p$ for O. We generated a slab model of the $p\,(4 \times 4)$ Ag(111) surface with three Ag layers and ≈ 18 Å of vacuum space (see Fig. 6.1). For structural optimizations of the slab model, we fixed the bottom layer and used total energy and force convergence thresholds of 0.01 eV/slab and 0.1 eV/Å, respectively. We sampled the Brillouin zone using a $3 \times 3 \times 1$, Γ-centered k-point grid. We also applied a dipole correction along (001) to cancel the artificial electric field across the slab. (180)

Random forests (RF) were trained using the scikit-learn package (version 0.19.1) for Python (version 3.6.5). (154) Processed data and Python scripts for the machine learning can be found in the SI. We removed highly correlated and near zero-variance descriptors from our data set. We randomly split the data set into a training and testing set with 2/3 and 1/3 of the data, respectively, so that we could estimate the out-of-sample error in the surface energy prediction.

6.3. Results

6.3.1. Surface Phase Diagram of Ag(111)

We perform a series of GCMC simulations, starting from the clean Ag(111) surface, under the conditions described above. Each chemical potential is simulated three times to improve the sampling of surface (composition and structure) phase space.

Fig. 6.2 shows the surface phase diagram generated by GCMC. There are three main regions of this phase diagram with respect to $\Delta\mu_O$ (see thick dotted lines). For $\Delta\mu_O \leq -0.51$ eV, the clean Ag surface is stable (see red line and Fig. 6.3A). From -0.51 eV $\leq \Delta\mu_O \leq -0.19$ eV, surface oxides form. At $\Delta\mu_O = -0.19$ eV, Ag undergoes a bulk phase transition to Ag_2O. Over 6000 structures were sampled by the GCMC simulations, and lines showing their surface free energy *vs.* $\Delta\mu_O$ are shown in gray. Practically speaking, each of the gray lines corresponds to an explicit DFT calculation of $U_{\text{slab}}^{\text{DFT}}$ in Eq. 6.10. We obtained a broad distribution of surface free energies, with values well below and above that of Ag(111) (see red line).

Four different structures make up the surface energy convex hull (see green line). For $\Delta\mu_O \leq -0.51$ eV, Ag(111) is preferred. Between -0.51 eV $\leq \Delta\mu_O \leq -0.49$ eV, one O per surface unit cell adsorbs onto an Ag_3-hollow site (see Fig. 6.3B). At $\Delta\mu_O$ above -0.49 eV and below -0.37 eV, surfaces oxides grow in the form of Ag_3O_4 pyramids (see Fig. 6.3C). O atoms at the corners of these pyramids bind to the surface at Ag_3-hollow sites. Under O-rich conditions, *i.e.* $\Delta\mu_O \geq -0.37$ eV, a continuous surface oxide layer forms with the composition $Ag_{10}O_7$ (see Fig. 6.3D). This surface consists of edge-sharing, distorted Ag_3O_4 and symmetric Ag_4O_5 pyramids. There is also an O atom at one of the two exposed, sublayer Ag_3-hollow sites.

This phase diagram, generated automatically using GCMC, is in excellent agreement with the experimental and theoretical literature on Ag(111). (55; 56; 277–279; 306–309; 316–319) The Ag_3O_4 pyramid is common to many of the structures that have been proposed. (55; 277; 279; 306–308; 316; 317) These pyramids can arrange themselves in a variety of geometries, such as Ag_2O(111)-like hexagons and shamrocks. (55; 277; 306–308; 316; 317) The $Ag_{10}O_7$ surface we find is very similar to a $c(4 \times 8)$ reconstruction that has been synthesized and, to date, has the lowest

reported DFT surface energy. (279) The main difference is that this structure contains unconnected chains of edge-sharing Ag_3O_4 pyramids whereas, in our structure, the chains are connected, which induces 4.52 meV/Å^2 increase in the surface energy. In this study, we impose $p\,(4 \times 4)$ surface periodicity based on historical precedent. However, oxide adlayers with different periodicities have been reported in the literature. (279; 309; 320) If we had imposed a smaller surface unit cell, Ag_3O_4 pyramids could still form but may not dimerize, thereby precluding the growth of 2D, continuous surface oxides like the $Ag_{10}O_7$ surface. We find that four unit cells perfectly fit pyramid dimers and that multiples of four are necessary to form chains of pyramid dimers. Therefore, while surface periodicity can affect the ground state arrangement of Ag_3O_4 pyramids, $p\,(4 \times 4)$ is ideal for computational studies because it is the smallest surface unit cell that can host 2D, continuous overlayers. It is noteworthy that ab initio GCMC, given only a few inputs and without any prior knowledge of the system, is able to reproduce the important features of the Ag(111) surface phase diagram, which took many decades to decipher.

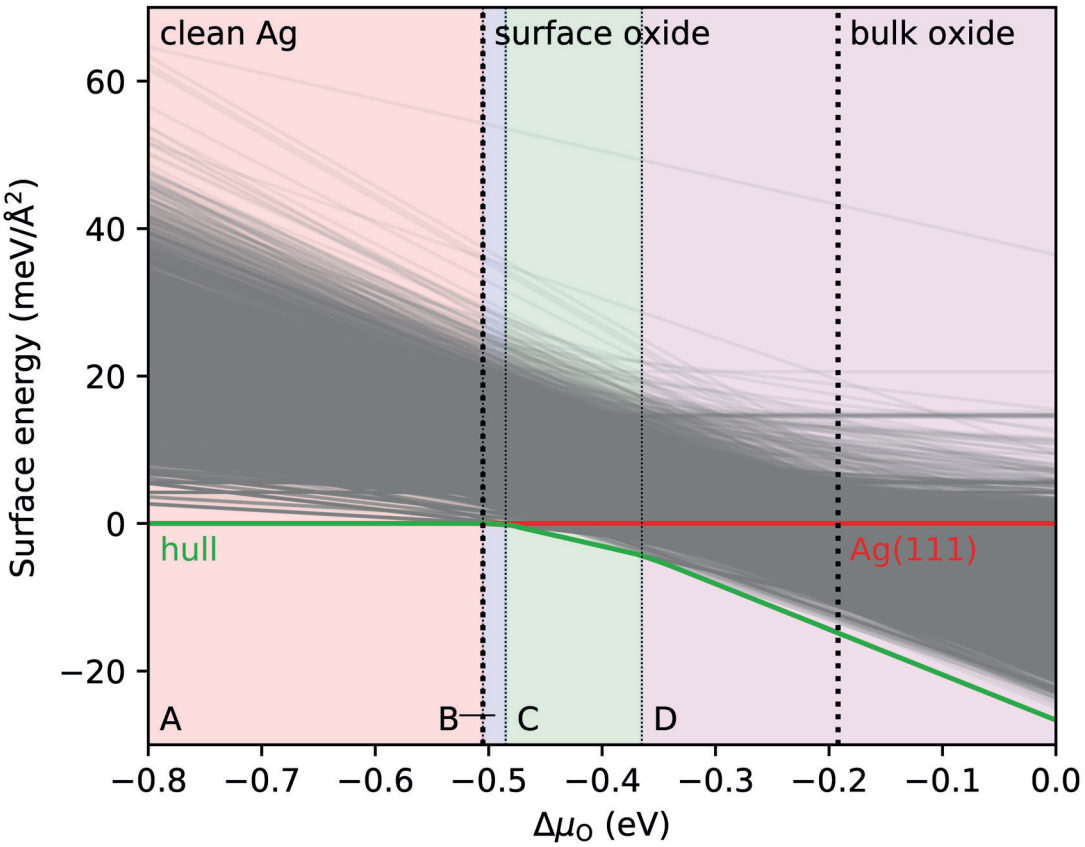

Figure 6.2: Surface phase diagram of Ag(111) exposed to O_2, generated by GCMC. There is a gray line for each surface sampled. The red and green lines correspond to Ag(111), *i.e.* the starting point, and the surface energy convex hull, respectively. Thick dotted lines separate the three main regions of the phase diagram, and thin dotted lines separate lightly shaded regions for the four surface phases (A-D, see Fig. 6.3) that constitute the hull.

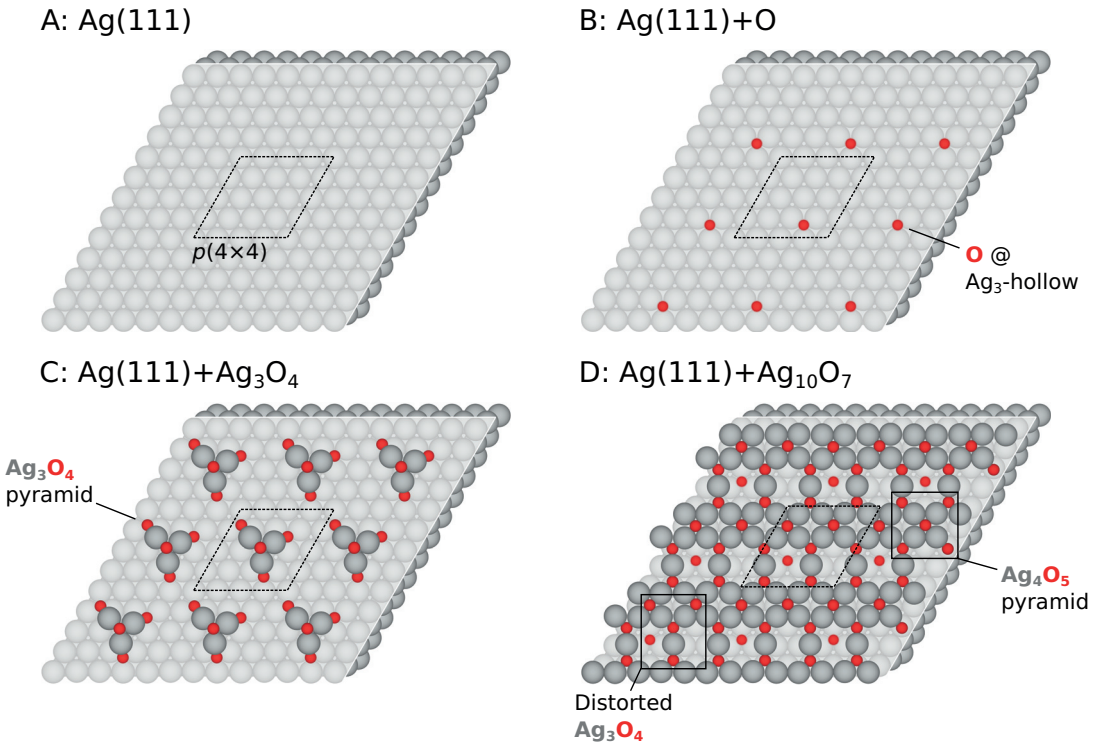

Figure 6.3: Stable Ag(111) surfaces and reconstructions discovered by GCMC: (A) clean Ag, (B) O at an Ag_3-hollow site, (C) formation of an Ag_3O_4 pyramid, and (D) growth of an $Ag_{10}O_7$ overlayer. All surfaces have $p\,(4 \times 4)$ periodicity.

6.3.2. Structural Descriptors for the Surface Energy

The GCMC simulations described in section 1 generate a large data set composed of structures and energies. This enables the use of machine learning, namely random forest (RF) regression, to determine the structural features that govern surface stability. We choose RF regression because we have shown previously that it is a powerful method for the discovery of structural and electronic descriptors for surface chemical properties like catalysis. (39) A RF is an ensemble of decision trees, each trained on a random subset of the data. The decision trees learn by splitting the data based on values of the independent variables (*e.g.* bond length = 2 Å) and then finding which of those splits best separates the data based on the dependent variable (*e.g.* surface energy). This type of learning is referred to as supervised because we know the value of the output for different sets of inputs. After supervised learning, the RF model can rank the importance of each feature and predict the surface energy (see Figs. 6.4-A and 6.4-B, respectively). Feature importance is a measure of how well splits based on each independent variable separate the data based on the dependent variable. We consider four types of structural features at the surface and calculate their averages: (1) bond length between atoms A and B ("bondAB"), (2) number of atom B within 3 Å of atom A ("cnAB"), (3) magnitude of the sum of the bond vectors (BV) pointing from atom A to all atom B within 3 Å ("bvAB"), and (4) z-component of the BV sum ("bvzAB"). Note that atoms A and B correspond to different elements. In addition to structural features, we calculate the global instability index ("gii"), which measures deviations of each atom from its preferred atomic valence. (321)

Figs. 6.4-A and 6.4-B show the importance of all the features and the goodness of fit of the RF model, respectively. The model has a root-mean-squared error (RMSE) of 2.16 meV/Å^2, and the data in Fig. 6.4-B lie very close to the perfect correlation

line. This result shows that we have included features that are excellent descriptors of the surface energy. Scatter plots of surface energy *vs.* the two most important descriptors, *i.e.* cnAgO and bvAgO (see Figs. 6.4-C and 6.4-D, respectively), reveal trends that help rationalize the stability of the $Ag_{10}O_7$ surface. Both plots have large spread in the surface energy and concave envelopes tracing the surface energy minima along the descriptor direction (see thin, black-dashed line). The sharpness of these envelopes near the surface energy minimum indicates that surfaces have a clear tendency for cnAgO = 2 and bvAgO $\approx$ 0.5. The former means that each surface Ag atom tends to form two bonds with O. The preferred value of bvAgO requires a more careful interpretation. bvAgO is zero when Ag either has no O neighbors or the Ag-O BV sums cancel. In the context of two-fold coordination of Ag with O, bvAgO is small when the O-Ag-O chain is slightly bent ($\approx 5^\circ$, see inset in Fig. 6.4D).

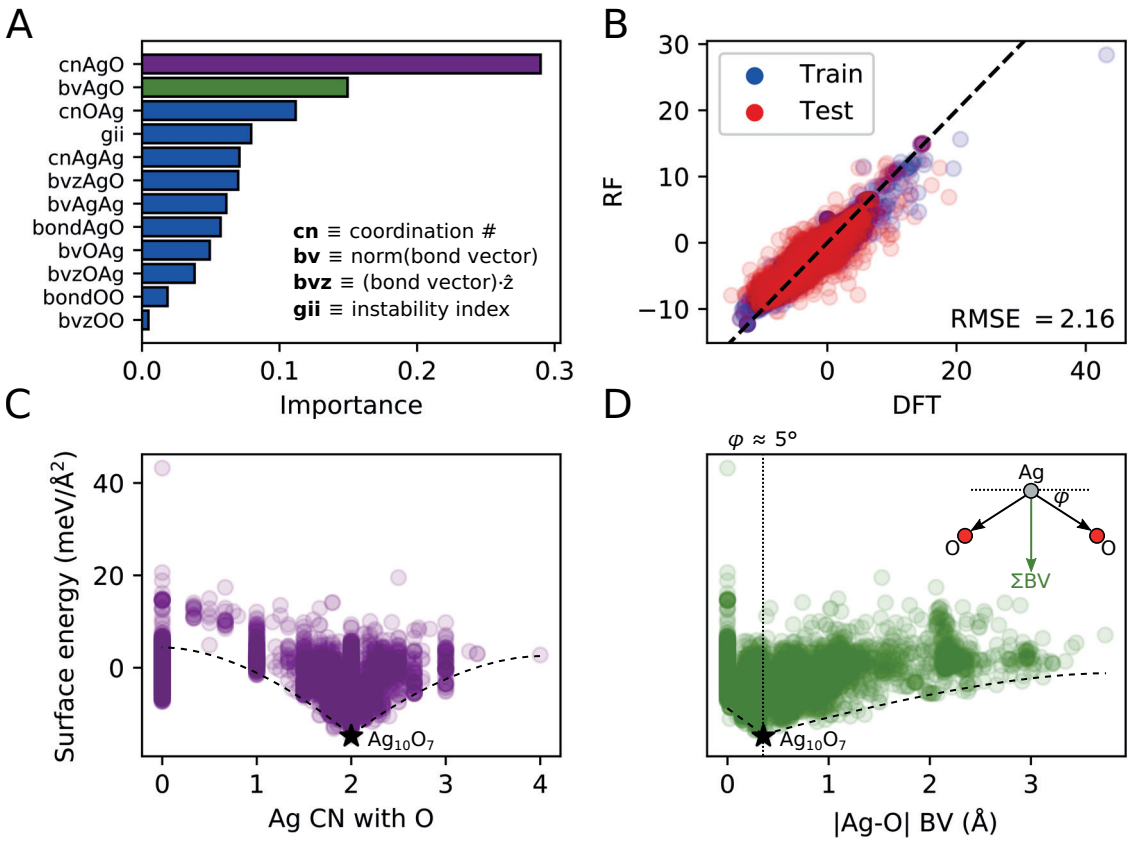

Figure 6.4: Analysis of structural descriptors for the surface energy. (A) Relative importance of descriptors calculated from random forest (RF) model. (B) Surface energy predicted by RF *vs.* DFT. Black-dashed line corresponds to perfect agreement. (C) Effect of the Ag coordination number (CN) with O (cnAgO) and (D) effect of the magnitude of Ag-O bond vector (BV) sum (bvAgO) on the surface energy. cnAgO is the number of O within 3 Å of Ag including bonding to the layer below. Thin, black-dashed lines highlight the trend of increasing free energy for deviations from ideal CN or BV sum. Stars denote the $Ag_{10}O_7$ surface.

6.3.3. Mechanistic Analysis of GCMC Composition-Structure Evolution Histories

While many oxide overlayers have been proposed, the mechanism of their formation remains unclear. (55; 56; 277–279; 306–309; 316–319) It is known that surface oxide formation requires facile O_2 dissociation and significant mass transport of Ag and O. (317; 318) A benefit of using GCMC is that it produces a composition-structure evolution history, which can be analyzed to reveal the stages of surface reconstruction. Fig. 6.5 shows the path taken by the GCMC simulation to obtain the $Ag_{10}O_7$ surface. There are three main stages of the mechanism: chain growth, pyramid formation, and pyramid dimerization. At the beginning of the first stage, pairs of O atoms adsorb onto nearby Ag_3-hollow sites (see Fig. 6.5-A). At the same time, they extract an Ag atom from the surface, forming surface O-Ag-O chains and subsurface Ag vacancies. Each single chain serves as a nucleation center from which longer chains can grow, such as double and branched chains, through the addition of extra Ag and O from their respective chemical potential reservoirs (see Figs. 6.5-B and 6.5-C, respectively). The latter is a critical intermediate in the formation of Ag_3O_4 pyramids in the second stage. Here, the branched chain reorients itself via O hopping between Ag_3-hollow and Ag_2-bridge sites (see Fig. 6.5-D) . After the pyramid forms (see Fig. 6.5-D) , the subsurface Ag vacancy is filled. Finally, in the last stage, pyramids dimerize. This starts with chain growth from one of the corners of the pyramid (see Fig. 6.5-E) . Once a double chain is formed (see Fig. 6.5-F) , it repositions itself (see Fig. 6.5-G) and, upon the deposition of Ag atom, forms a dimer (see Fig. 6.5-H) , which is the main repeating unit of the $Ag_{10}O_7$ surface.

Since many proposed overlayers express the pyramid motif, (55; 56; 277–279; 306–309; 316–319) this is a plausible mechanism for their formation as well, except for the third stage. Recall that the best descriptors of surface energy from the random forest

model are the Ag-O CN and BV sum. Not only do the most stable surfaces exhibit ideal values for these descriptors, but their building blocks, *i.e.* chains and pyramids, do as well. This shows that, within the context of this mechanism, Ag-O CN and BV sum are the key driving forces behind the reconstruction of Ag(111).

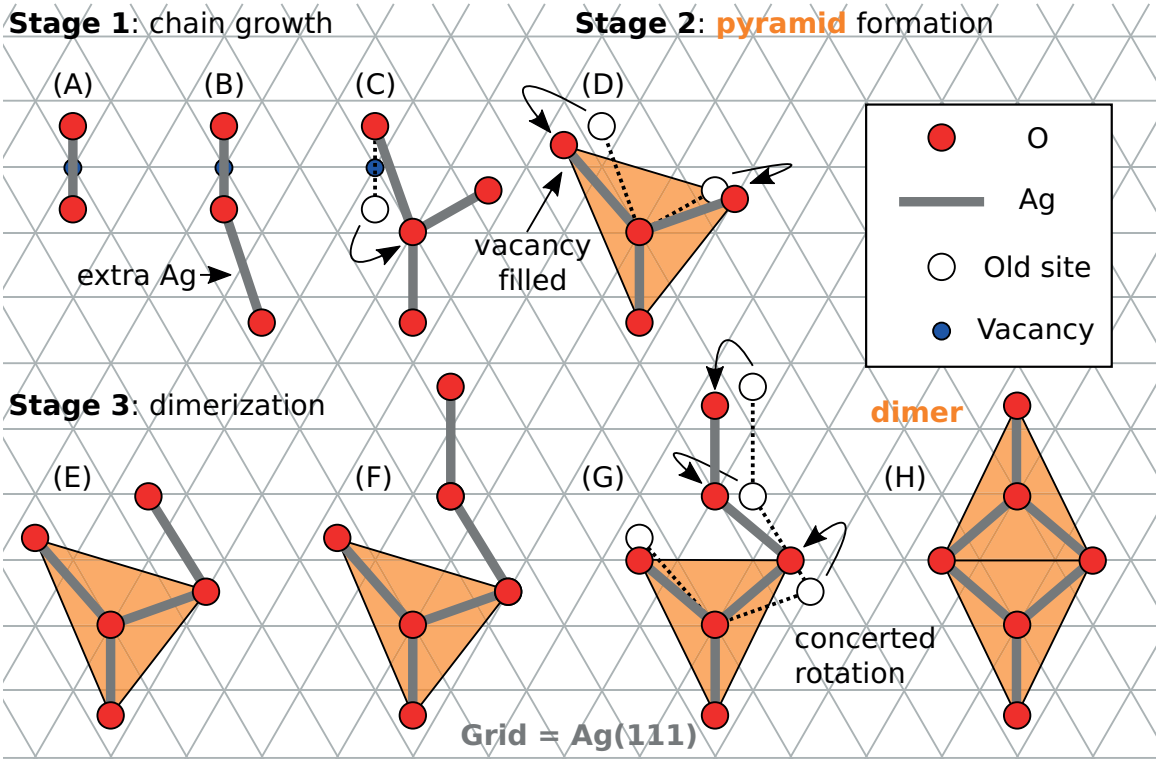

Figure 6.5: Mechanism for the formation of the $Ag_{10}O_7$ surface (see Fig. 6.3D). Red, white, and blue circles correspond to O atoms, their previous position, and subsurface Ag vacancies, respectively. Ag atoms are represented by a thick gray line. The mechanism involves three stages: (A-C) chain growth, (D) pyramid formation, and (E-H) dimerization. In the chain growth stage, (A) an O-Ag-O chain and subsurface Ag vacancy form followed by (B) linear and (C) branched chain growth. (D) Next, O atoms jump to new sites and the subsurface Ag vacancy is filled, forming an Ag_3O_4 pyramid. Finally, in the dimerization stage, (E-F) linear chains grow from the pyramid, which (G) undergoes a concerted rotation. Upon the deposition of a Ag atom, a pyramid dimer is formed.

6.4. Discussion

Here, we discuss the strengths and weaknesses of *ab initio* GCMC and provide some recommendations for its future application. Its strengths are that it requires few inputs and has minimal bias toward a particular solution. There are two parameters per element for the configurational bias (r_{min} and r_{max}), one parameter defining the dimensions of the variable composition region (see Fig. 6.1), and three parameters for the GCMC simulation (T, μ_{Ag}, and μ_{O}). In *ab initio* thermodynamics studies of surface reconstructions, it is common practice to generate a set of reasonable trial structures, sometimes numbering in the hundreds. (277; 279; 316–319) It is difficult, however, to remove bias from this procedure when structures are human-selected. Such biases are avoided in GCMC because each structure is selected proportionally to its weight in the grand canonical ensemble, which more closely resembles selection in nature. The weaknesses of *ab initio* GCMC are that it relies on costly *ab initio* calculations and only works for surfaces. The first weakness can be overcome by replacing DFT with reactive force fields (*e.g.* Reaxff, (322) REBO, (323; 324) and COMB (325)) or machine learning atomistic potentials (*e.g.* aenet (326; 327)). There is a trade-off, however, because these methods require careful parameterization and testing, and are only available for a small but growing set of systems. A current limitation of our software is that it can only be used for the study of surfaces. We are already in the process of generalizing the code so that it can also be used to study bulk materials and nanoparticles.

Given its success with surfaces, we believe that *ab initio* GCMC will become an important tool for research in heterogeneous catalysis. Take, for example, the epoxidation of ethylene over Ag, where it is believed that an electrophilic surface O species, seen in XPS measurements, is the key to selective formation of ethylene

oxide. (316; 317; 328–336) For the reconstructions of Ag(111), however, this species is not observed, thus leading to the conclusion that the stable surfaces are not responsible for catalysis. (333) Since GCMC samples both stable and unstable structures, it may find surfaces that do possess electrophilic O species and can therefore catalyze selective epoxidation. In practice, this could involve three steps: (1) reach equilibrium with μ_{Ag} and μ_{O}, (2) introduce ethylene chemical potential reservoir $\mu_{\mathrm{C_2H_4}}$, and (3) reach equilibrium with μ_{Ag}, μ_{O}, and $\mu_{\mathrm{C_2H_4}}$. Alternatively, we could apply a bias toward higher free energies that would increase the sampling of surfaces that are less stable but potentially more catalytically active. Other promising applications of *ab initio* GCMC include the study of binary and ternary materials (*e.g.* TiO_2 and $SrTiO_3$), reentrant transitions, solvation by including a solvent chemical potential reservoir, and nanoparticle growth for crystal structure prediction.

6.5. Conclusions

In this chapter, we introduce our new method of *ab initio* GCMC for the investigation of surface reconstruction. This method requires a minimal number of selected parameters, enables surfaces to evolve under realistic conditions, and reduces bias associated with selection of trial structures for surface stability analyses. We show that *ab initio* GCMC reproduces the salient features of the Ag(111) surface phase diagram, which took decades to unravel, and, in particular, finds a surface ($Ag_{10}O_7$) that is in excellent agreement with the most stable surface reported in the literature. By analyzing the composition-structure evolution histories of GCMC simulations, we propose a mechanism, based on O-Ag-O chain growth and rearrangement, that can explain the formation of Ag_3O_4 pyramid building blocks, which are common to a number of nearly-stable reconstructions of Ag(111). We also show the advantages of using GCMC to generate data for the discovery of structural descriptors of the surface

energy via machine learning. We find that the most relevant descriptors (coordination number of Ag with O and norm of the Ag-O BV sum) support our proposed mechanism and therefore are key driving forces for reconstruction. *Ab initio* GCMC, from structure generation to analysis, is fully transferable to the study of the surfaces of other materials and also holds promise for the exploration of other processes, such as heterogeneous catalysis and nanoparticle growth.

APPENDIX A : Supplemental: Phosphorus-decorated reconstructions of the (0001) surface of Ni_2P and Ni_5P_4

Table A.1: Calculated and experimental bulk parameters in Å

	$\mathbf{Ni_2P}$ (181)		$\mathbf{Ni_5P_4}$ (337)	
	Calc.	*Exp.*	*Calc.*	*Exp.*
$a = b$	5.86	5.87	6.78	6.79
c	3.35	3.39	10.97	10.99

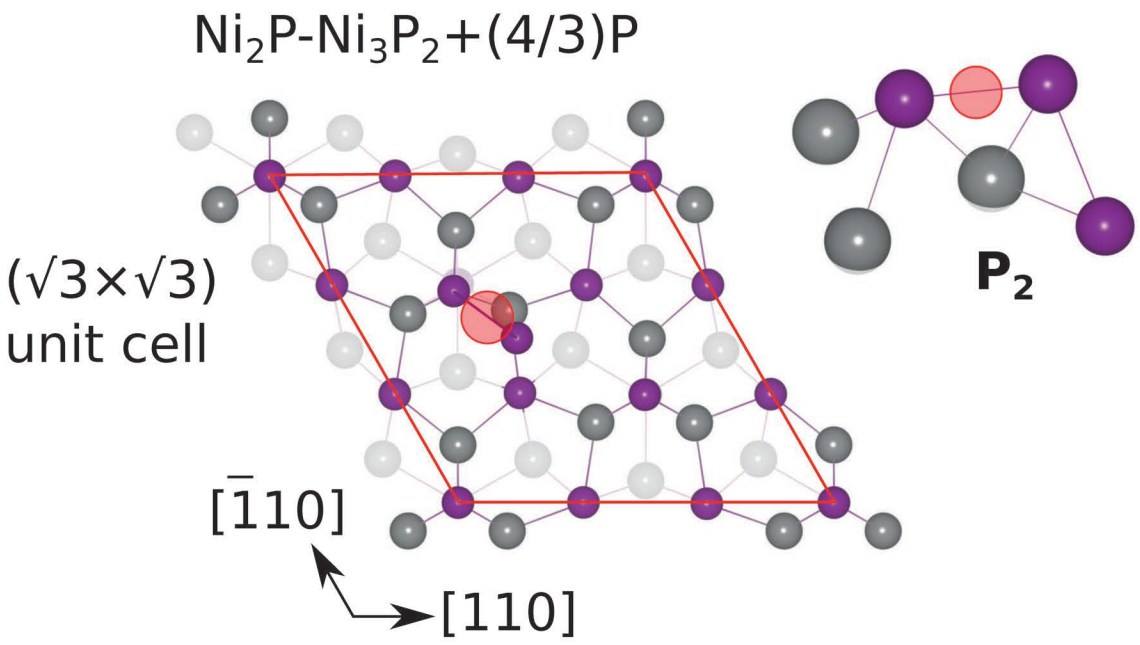

Figure A.1: Surface crystal structure of $Ni_2P(0001)$ with a Ni_3P_2+(4/3)P termination. Structural inset highlights the P_2 complex referenced in Chapter 3. Red lines outline the $\sqrt{3} \times \sqrt{3}$ $R30°$ supercells.

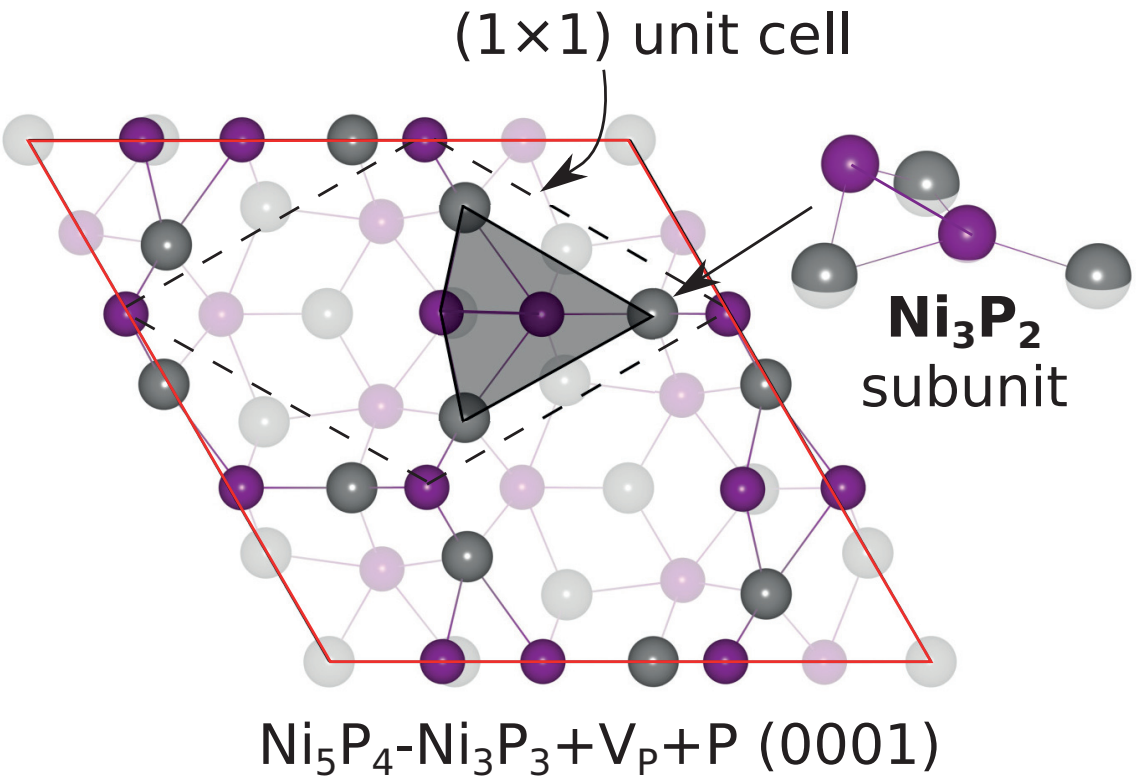

Figure A.2: Surface crystal structure for the $Ni_3P_3+V_P+P$ reconstruction of $Ni_5P_4(0001)$. Shaded region corresponds to the inset highlighting the Ni_3P_2 referenced in Chapter 3.

A.1. Survey of the $Ni_5P_4(000\bar{1})$ Layers and Reconstructions

Here, the Ni_3P_2 and Ni_3P_3-derived reconstructions of $Ni_5P_4(000\bar{1})$ surfaces (stacking shown in Fig. A.3a) are discussed. The discussion on the Ni_4P_3-derived termination and it's reconstructions is found in Chapter 3.

A.1.1. Ni_3P_2-derived Surfaces of $Ni_5P_4(000\bar{1})$

Fig. A.3b shows the $\sqrt{3} \times \sqrt{3}$ $R30°$ bulk-derived Ni_5P_4-Ni_3P_2 layer. This layer is composed of repeated, corner-sharing small and big Ni_3P subunits. The small Ni_3P subunit has a trigonal pyramidal geometry with P pointing out of the surface plane. In the big Ni_3P subunit, however, P is coplanar with the three Ni atoms. The surface structure of Ni_5P_4-$Ni_3P_2(000\bar{1})$ is identical to that of Ni_5P_4-Ni_3P_3+$2V_P$+P(0001), which shares the same layer composition. The removal of P from the small Ni_3P subunit is unfavorable and destabilizes the surface (see surface phase diagram in Fig. A.4a, compare solid line, Ni_3P_2, and dotted line, Ni_3P_2+V_P). The adsorption of P at the small Ni_3 subunit and the center of the hollow region is also unfavorable (structure not shown), but less so than the formation of P vacancies, therefore leading to a less significant destabilization of the Ni_5P_4-Ni_3P_2 surface.

A.1.2. Ni_3P_3-derived Surfaces of $Ni_5P_4(000\bar{1})$

Fig. A.3c shows the $\sqrt{3} \times \sqrt{3}$ $R30°$ Ni_5P_4-Ni_3P_3 layer. This layer is composed of closed-chains Ni_3P_3 subunits. The Ni_3P_3 subunit is characterized by a triangular Ni_3 overlaid on a P_3 that are rotated 60° relative to one another. Ni corners of different Ni_3P_3 subunits converge at single point (translucent red circle in Fig. A.3c) thus forming a Ni_3-hollow site. The same is true for the P corners, which form a trinuclear P_3 site. The formation of vacancies, both Ni and P, are unfavorable

compared to the saturation of each Ni_3-hollow site with three P. The additional P results to staggered Ni_3P_3 subunits, which stabilizes the surface by $\approx$ 0.5-1 J/m^2. Neither the bulk-derived Ni_5P_4-Ni_3P_2 (discussed in the previous section) nor the reconstructed Ni_5P_4-Ni_3P_3+3P surface is more stable than Ni_5P_4-Ni_4P_3+3P (shown in Fig. 3.3d in Chapter 3).

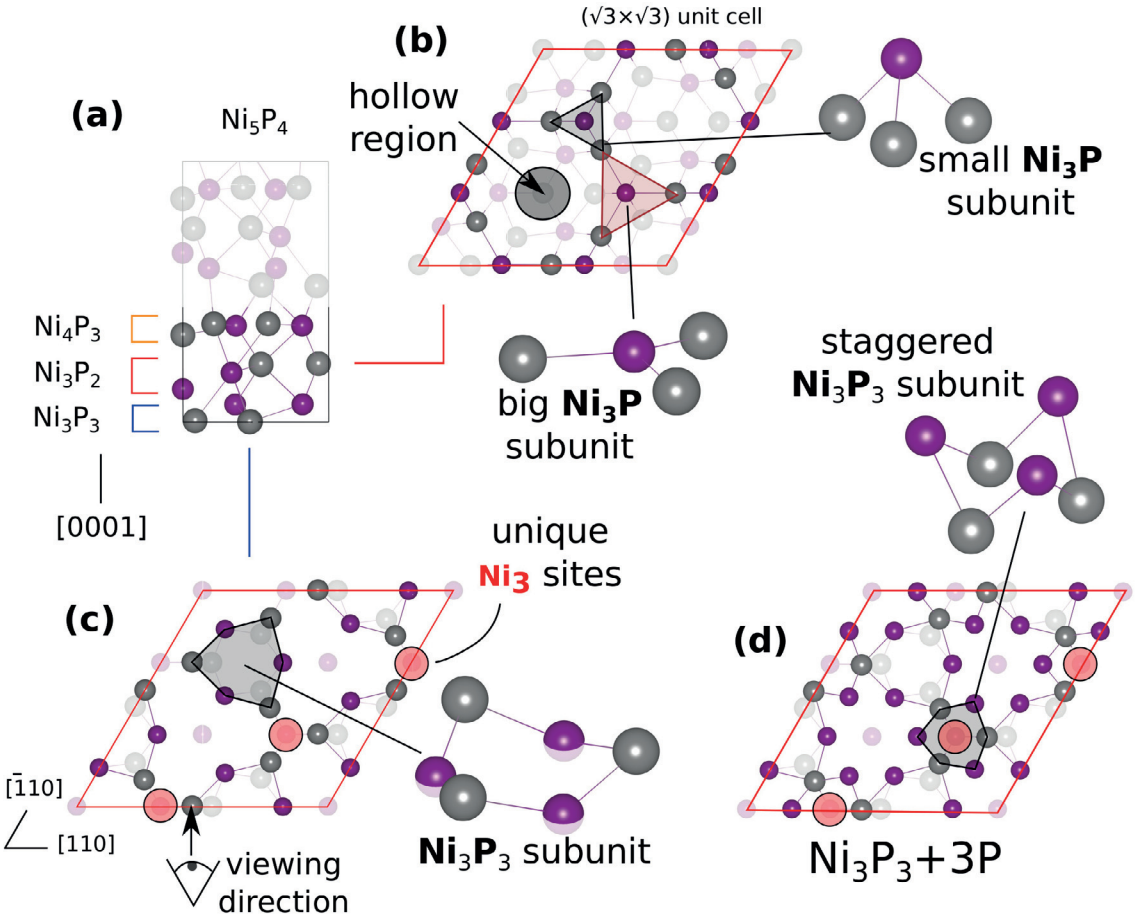

Figure A.3: Surface crystal structure for bulk-derived terminations and reconstructions of $Ni_5P_4(000\bar{1})$. (a) Bulk layering in Ni_5P_4. Bulk-like $(000\bar{1})$ terminations Ni_3P_2 (b) and Ni_3P_3 (c) with shaded regions corresponding to the insets highlighting important structural features. (d) Stable $(000\bar{1})$ reconstruction Ni_3P_3+3P.

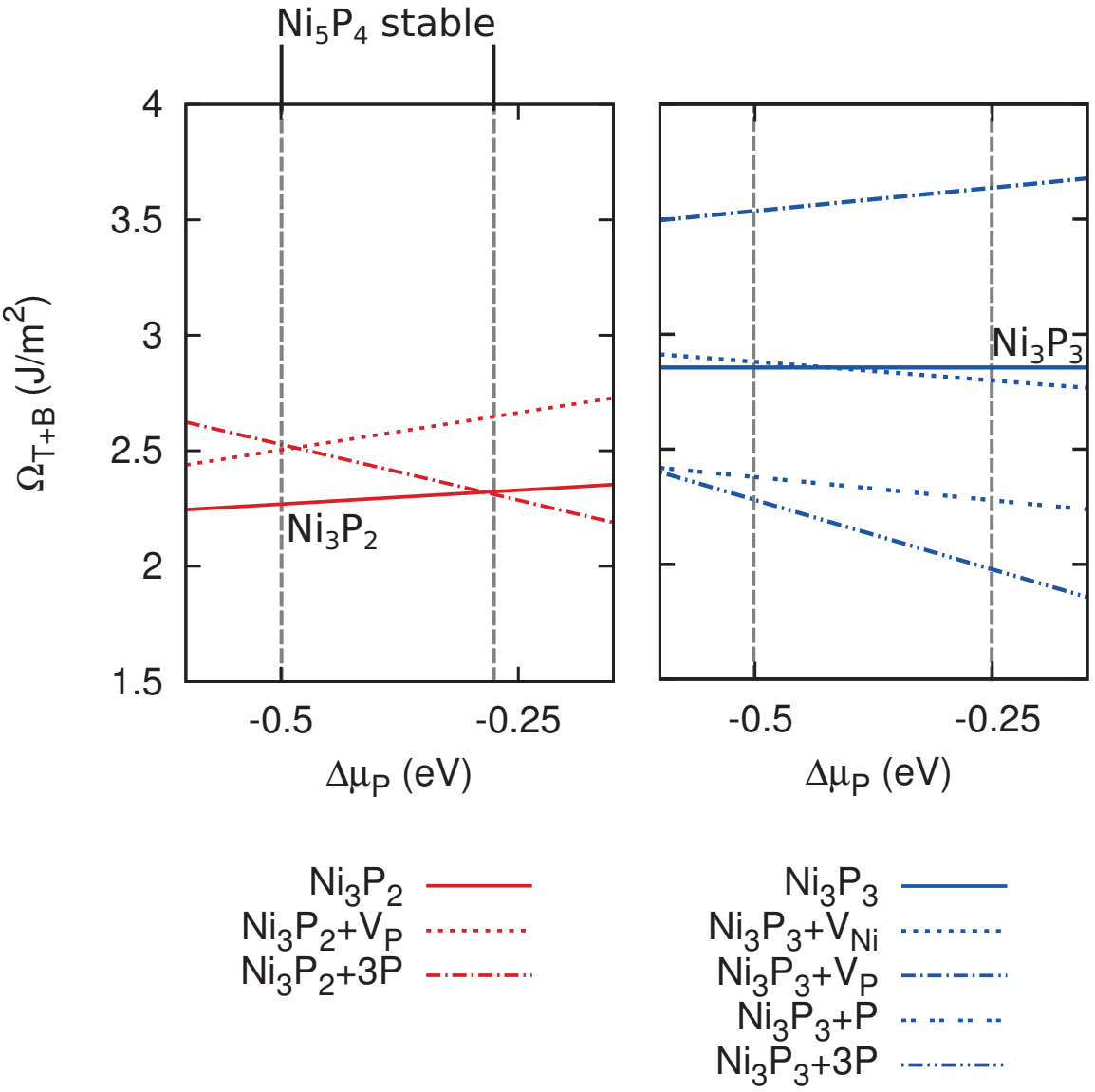

Figure A.4: Surface phase diagram for $Ni_5P_4(000\bar{1})$ surfaces as a function $\Delta\mu_P$ (eV). Ω_{T+B} corresponds to the combined surface energies (J/m^2) of the top (T) and bottom (B) surfaces of (a) Ni_3P_2 and (b) Ni_3P_3 . There are no regions of $\Delta\mu_P$ where these bulk-terminations or reconstructions are favored. See also Fig. 3.4d in Chapter 3 for the corresponding phase diagram of the Ni_4P_3-derived surfaces.

Table A.2: Calculated Löwdin charges for the surface atoms for the stable surface reconstructions and some of the bulk-derived terminations of Ni_2P and Ni_5P_4(0001) and $(000\bar{1})$.

Composition		**Species**	**Location**	**No. Electrons**			**Net Charge**
Bulk	*Surface*			*s*	*p*	*d*	[*e*]
Ni_2P	Ni_3P	P	Ni_3P subunit	1.63	3.34	0.76	-0.72
		Ni	In-plane	0.64	5.98	8.74	0.65
	Ni_3P_2	P	Ni_3P subunit	1.61	3.23	0.82	-0.66
		Ni	In-plane	0.72	5.97	8.66	0.64
	Ni_3P_2 +P	P	Ni_3P subunit	1.60	3.19	0.88	-0.68
			Ni_3-hollow	1.75	3.09	0.46	-0.30
		Ni	In-plane	0.73	5.97	8.55	0.75
Ni_5P_4 (0001)	Ni_3P_2	P	Ni_3P subunit	1.61	3.35	0.73	-0.68
			P subunit	1.41	2.84	0.76	-0.01
		Ni	In-plane	0.65	5.98	8.72	0.65
	Ni_3P_3 $+V_P+P$	P	P_2	1.62	2.95	0.62	-0.20
			Big Ni_3P subunit	1.44	3.14	1.05	-0.64
			Small Ni_3P subunit	1.73	3.14	0.46	-0.33
		Ni	In-plane	0.69	5.97	8.58	0.76
	Ni_3P_3 $+2V_P+P$	P	Big Ni_3P subunit	1.62	3.32	0.73	-0.68
			Small Ni_3P subunit	1.74	3.19	0.42	-0.34
		Ni	In-plane	0.73	5.97	8.58	0.72
	Ni_4P_3	P	Ni_4P_3 subunit	1.62	3.19	0.80	-0.62
		Ni	Ni_4P_3 satellite	0.68	5.98	8.72	0.62
			Ni_4P_3 center	0.69	5.97	8.55	0.79
	Ni_4P_3 +(8/3)P	P	Ni_4P_3 subunit	1.46	3.07	1.08	-0.61
			Ni_3-hollow	1.61	3.06	0.65	-0.33
			P_3-hollow	1.55	2.78	0.59	0.08
			P_2	1.62	2.94	0.51	-0.06
		Ni	In-plane	0.66	5.98	8.62	0.74
	Ni_4P_3 +3P	P	Ni_4P_3 subunit	1.46	3.06	1.09	-0.61
			Ni_3-hollow	1.55	3.05	0.72	-0.32
			P_3-hollow	1.55	2.77	0.61	0.07
			P_2	1.63	2.93	0.49	-0.05
		Ni	In-plane	0.65	5.98	8.63	0.74
Ni_5P_4 $(000\bar{1})$	Ni_3P_3	P	In-plane	1.52	3.09	0.81	-0.41
		Ni		0.72	5.98	8.68	0.62
	Ni_4P_3	P	Ni_4P_3 subunit	1.62	3.18	0.82	-0.61
		Ni	In-plane	0.69	5.98	8.68	0.66
	Ni_4P_3 +3P	P	Ni_4P_3 subunit	1.46	3.05	1.08	-0.59
			Ni_3-hollow	1.55	3.10	0.70	-0.36
			P_3-hollow	1.53	2.78	0.61	0.07
			P_2	1.60	2.92	0.58	-0.11
		Ni	Ni_4P_3 satellite	0.67	5.98	8.60	0.75
			Ni_4P_3 center	0.66	5.98	8.66	0.70

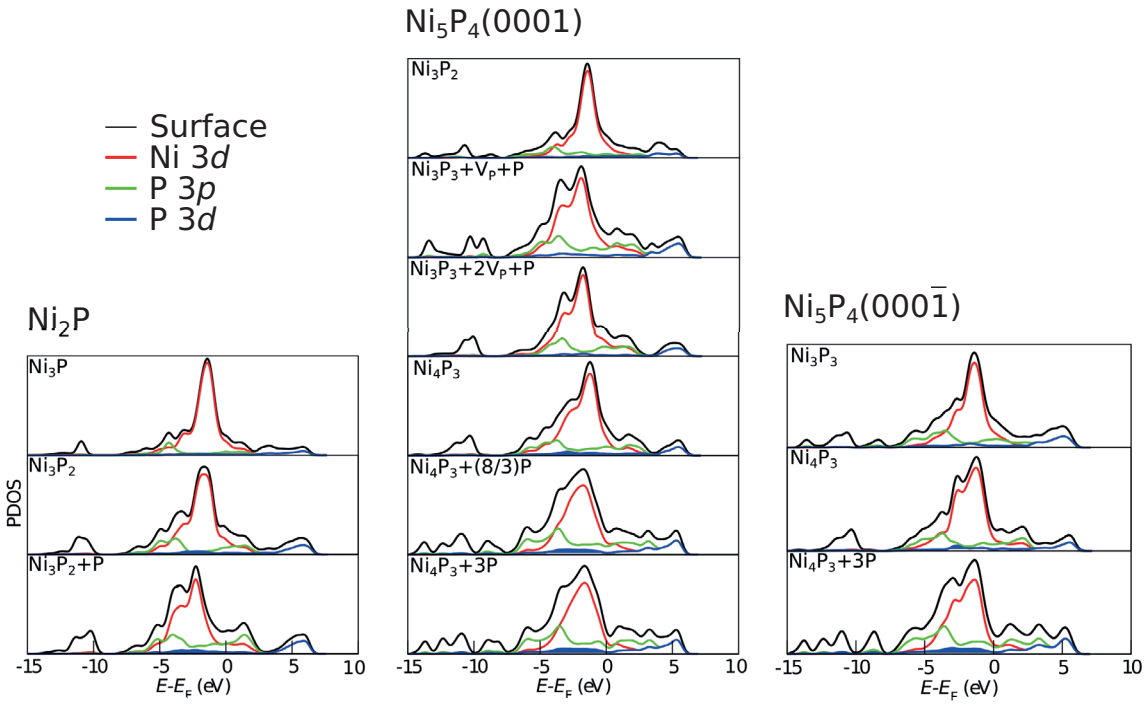

Figure A.5: Orbital-projected density of states (PDOS) for the surface atoms for the stable surface reconstructions and some of the bulk-derived terminations of Ni_2P and $Ni_5P_4(0001)$ and $(000\bar{1})$.

APPENDIX B : Supplemental: Mechanism of H_2 evolution on aqueous reconstructions of the (0001) surface of Ni_2P and Ni_5P_4: the crucial role of phosphorus

B.1. Additional Computational Details

For geometry relaxations, we used the Broyden-Fletcher-Goldfarb-Shanno (BFGS) quasi-newton algorithm, based on the trust radius procedure. To speed up the self-consistent field (SCF) procedure, electron charge densities were mixed using local-density-dependent Thomas-Fermi screening with a mixing factor of 0.7. For vibrational mode calculations, we choose a mixing factor of 0.1 for updating the SCF potential.

For Ni, there are four orbitals in our pseudopotential: $3s$, $3p$, $4s$, and $3d$. The cutoff radius for these pseudo orbitals were 1.80, 1.86, 1.82, and 1.98 Bohr, respectively. Our pseudopotential for P was derived from explicitly pseudizing $3s$, $3p$, and $3d$ orbitals with cutoff radii of 1.50, 1.65, and 1.95 Bohr, respectively. The $3d$ orbital was included to permit the formation of anionic and hypervalent P species. For H, the $1s$ orbital was pseudized with a cutoff radius of 0.77 Bohr. A step function augmentation operator (105) was applied was applied with a step-height of -1.60 Ha and step-width from 0.10 to 1.60 Bohr using the $3s$ as the local potential. This is done to strategically target the $3d$ orbital of Ni and improve the transferability of the pseudopotential.

B.2. Additional Theoretical Details

We define the first term in Eq. 4.8 as the desorption free energy for A to form A(std)

$$\Delta G_{\mathrm{A,diss}} = \Delta G_{\mathrm{dsrp}} + (G_{\mathrm{H}_x\mathrm{AO}_y^z} + n_{\mathrm{H}} G_{\mathrm{H}} + n_e G_e - G_{\mathrm{A(std)}} - n_w G_{\mathrm{H_2O}}) \tag{B.1}$$

We approximate the free energies in ΔG_{dsrp} (first term in Eq. 4.8) with DFT total energies (338). For surface adsorbates, we compute corrections for the zero-point energy (ZPE). For gas phase species, we use the experimental ZPE, integrated heat capacity ($H - H^\circ$), and standard molar entropy (S^0). (339) The resulting expression for ΔG_{dsrp} is

$$\Delta G_{\mathrm{dsrp}} = \Delta E_{\mathrm{dsrp}}^{\mathrm{DFT}} + \Delta \mathrm{ZPE} + \Delta(H - H^\circ) - T\Delta S^\circ \tag{B.2}$$

where T is the temperature in K. The terms inside the second bracket in Eq. 4.8 can be re-expressed in the following way:

$$\begin{aligned} \Delta G_{\mathrm{A,diss}} = \Delta G_{\mathrm{dsrp}} + (G^\circ_{\mathrm{H}_x\mathrm{AO}_y^z} + n_{\mathrm{H}} G^\circ_{\mathrm{H}} + n_e G^\circ_e - G^\circ_{\mathrm{A(std)}} - n_w G^\circ_{\mathrm{H_2O}}) \\ + k_{\mathrm{B}} T \ln a_{\mathrm{H}_x\mathrm{AO}_y^z} - 2.303 n_{\mathrm{H}} k_{\mathrm{B}} T p\mathrm{H} - n_e q_e U \end{aligned} \tag{B.3}$$

where G° is the standard free energy, k_{B} is Boltzmann's constant, a is the activity, and q_e is the electron charge. We assume an ideal solution, which allows us to replace activity with concentration. The standard oxidation free energy (or reduction free energy for $n_e < 0$) of A(std) to form $\mathrm{H}_x\mathrm{AO}_y^z$ in acidic medium *vs.* SHE is defined as

$$\Delta G^\circ_{\mathrm{A(std)/H}_x\mathrm{AO}_y^z} = G^\circ_{\mathrm{H}_x\mathrm{AO}_y^z} + n_{\mathrm{H}} G^\circ_{\mathrm{H}} + n_e G^\circ_e - G^\circ_{\mathrm{A(std)}} - n_w G^\circ_{\mathrm{H_2O}} \tag{B.4}$$

The experimental $\Delta G^{\circ}_{\mathrm{A(std)}/\mathrm{H}_x\mathrm{AO}_y^z}$ of a number of solid and aqueous Ni and P species can be found in Table B.1.

Table B.1: Experimental standard formation free energy of solid and aqueous Ni and P species. (1) n corresponds to the stoichiometric coefficients in Eq. B.4.

A	$H_xAO_y^{z-}$	$\mu^0_{H_xAO_y^{z-}}$ kcal/mol	eV	n_A	n_{H_2O}	n_{H^+}	n_{e^-}
Ni	Ni(s)	0	0	1	0	0	0
	Ni^{2+}(aq)	-10.9	-0.473	1	0	0	2
	NiO(s)	-50.6	-2.19	1	1	2	2
	$NiOH^+$(aq)	-54.4	-2.36	1	1	1	2
	$Ni(OH)_2$(s)	-106.9	-4.636	1	2	2	2
	$Ni(OH)_2$(aq)	-86.1	-3.73	1	2	2	2
P	P(s)	0	0	1	0	0	0
	P_2(g)	24.8	1.08	2	0	0	0
	P_4(g)	5.85	0.254	4	0	0	0
	PO_4^{3-}(aq)	-244.0	-10.58	1	4	8	5
	$P_2O_7^{4-}$(aq)	-459.8	-19.94	2	7	14	10
	P_4O_{10}(s)	-644.8	-27.96	4	10	20	20
	PH_3(g)	3.2	0.14	1	0	-3	-3
	PH_3(aq)	0.35	0.015	1	0	-3	-3
	PH_4^+(aq)	16.2	0.702	1	0	-4	-3
	HPO_4^{2-}(aq)	-260.91	-11.314	1	4	7	5
	$H_2PO_4^-$(aq)	-270.73	-11.740	1	4	6	5
	H_3PO_4(s)	-267.5	-11.60	1	4	5	5
	H_3PO_4(aq)	-273.10	-11.843	1	4	5	5
	$H_3PO_4 \cdot 0.5H_2O$(s)	-296.9	-12.87	1	4.5	5	5
	PH_4OH(aq)	-56.34	-2.443	1	1	-3	-3
	$HP_2O_7^{3-}$(aq)	-472.5	-20.49	2	7	13	10
	$H_2P_2O_7^{2-}$(aq)	-481.6	-20.88	2	7	12	10
	$H_3P_2O_7^-$(aq)	-484.7	-21.02	2	7	11	10
	$H_4P_2O_7$(aq)	-486.8	-21.11	2	7	10	10

B.3. Aqueous Stability of Bulk Ni_2P and Ni_5P_4

A precondition for the existence of a surface is existence of a stable bulk phase that supports it. Under UHV conditions with chemical potential reservoirs for Ni and P, we previously defined the regions of μ_{Ni} and μ_{P} where different bulk nickel phosphide compositions (Ni_xP_y) are stable. (67) Under electrochemical conditions, however, we consider the equilibrium between bulk nickel phosphides and aqueous ions of Ni and P, *e.g.* Ni^{2+}(aq) and H_3PO_4(aq) and some of their solid phases. Depending on the pH or U, aqueous ions of Ni and P with different charge states are possible. Fig. B.1 shows the Pourbaix diagram for Ni and P at 1 M and 300 K. Under reducing conditions, face-centered cubic Ni(s), oxidation state (o.s.) = 0, and PH_3, P o.s. = -3, are favorable. In acidic and oxidizing conditions, Ni prefers an o.s. of +2 and forms a coordination complex with six water molecules. At a pH of 6, NiO becomes the predominant phase while maintaining the same o.s. At oxidizing potentials, P^{5+} goes through a progression from phosphoric acid (H_3PO_4) at low pH to phosphate (PO_4^{3-}) at high pH.

Since Ni (P) has a small positive (negative) charge in nickel phosphides, (67) we expect that bulk Ni_2P and Ni_5P_4 are stable along the equilibrium line bewteen neutral and oxidized Ni and P species. The bulk stability regions of Ni_2P and Ni_5P_4 at 300 K are plotted in Fig. B.2. As the concentration of aqueous Ni and P decreases, with the constraint that $\beta[\mathrm{NiO}_a^b] = \alpha[\mathrm{H}_x\mathrm{PO}_y^z]$ where α and β are the stoichiometric coefficients for Ni and P in the bulk Ni phosphide, naturally so does the size of the bulk stability region. Above the bulk stability regions, Ni_xP_y follows oxidative degradation pathways forming either Ni^{2+} and $H_3PO_4/H_2PO_4^-$ or NiO and HPO_4^{2-}/PO_4^{3-}. Below, Ni_xP_y is reduced to Ni(s) and PH_3(g). For a 1 M concentration of solvated species, the bulk stability region of Ni_2P is larger than that of Ni_5P_4 indicating that Ni_5P_4

has a more limited operational potential in an aqueous electrochemical environment than Ni_2P, especially under more reducing conditions ($U <$ -0.50 V). We compare the stability of aqueous surface reconstructions only within the bulk stability region.

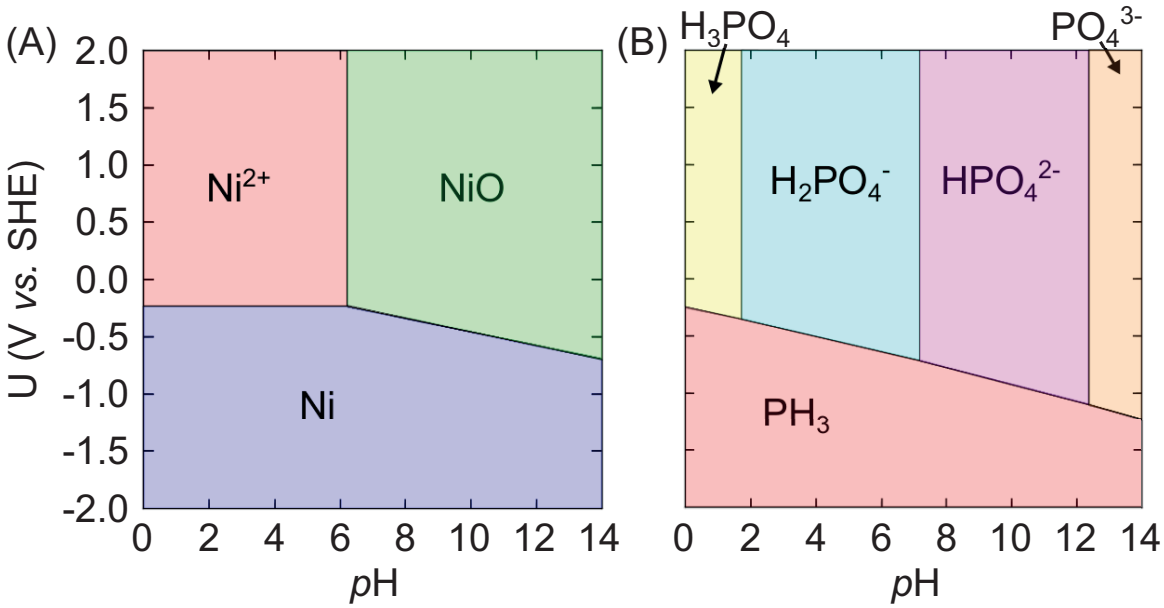

Figure B.1: Experimental Pourbaix diagram for a 1 M aqueous solution of (A) Ni and (B) P at 298.15 K constructed using data shown in Table B.1.

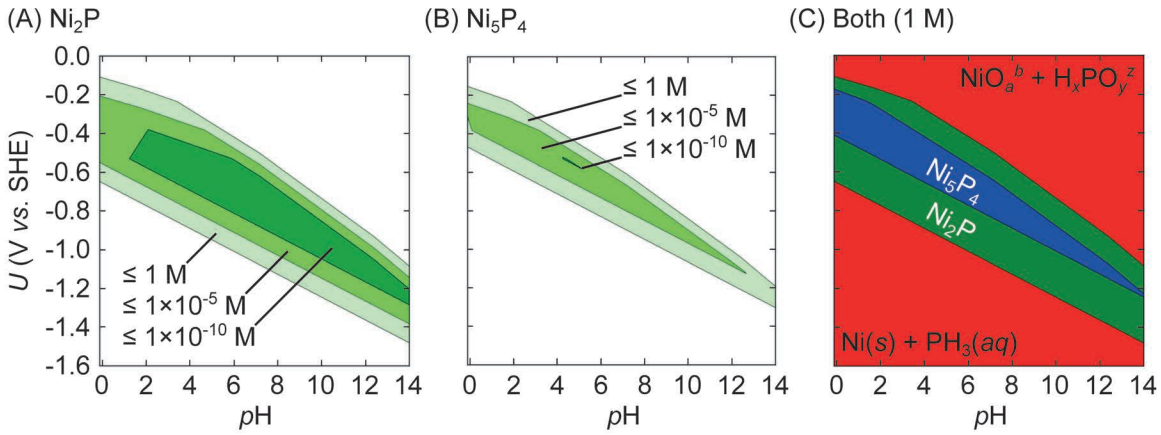

Figure B.2: Bulk phase diagram of (A) Ni_2P, (B) Ni_5P_4 for different molar concentrations of solvated species and (C) Ni_xP_y for 1 M NiO_a^b at 300 K. Refer to Fig. B.1 for the values of a, b, x, y, and z at a specific U and pH.

B.4. How To Use Tables B.2-B.4

We use the data in Tables B.2-B.4 to calculate the relative stability between surface phases (see Figs. 4.1 and 4.2). For a demonstration, consider the phase $Ni_2P(s)/Ni_3P_2(0001)+(5/3)H$ (red region) shown in Fig. 4.1. We calculate the free energy of surface phases relative to a reference surface, *e.g.* for $Ni_2P(0001)$, Ni_3P_2+P, at 1 M $[H_xPO_y^z]$ (2 M $[H_xNiO_y^z]$) and 298.15 K.

For $p\mathrm{H} < 1$, $U < -0.31$ V, and a 1×1 surface unit cell, the reaction to form phase 1 from the reference surface is

$$[Ni_3P_2 + P]^{\mathrm{ref}} + (14/3)(H^+ + e^-) \rightleftharpoons [Ni_3P_2 + (5/3)H] + PH_3(aq) \tag{B.5}$$

where the elements enclosed in brackets correspond to a single surface. To work with whole numbers instead of fractions, we consider a $\sqrt{3} \times \sqrt{3},\ R30°$ surface unit cell for which Eq. B.5 becomes

$$[Ni_9P_6 + 3P]^{\mathrm{ref}} + 14(H^+ + e^-) \rightleftharpoons [Ni_9P_6 + 5H] + 3PH_3(aq) \tag{B.6}$$

We can split this reaction into two steps

$$[Ni_9P_6 + 3P]^{\mathrm{ref}} + 5(H^+ + e^-) \rightleftharpoons [Ni_9P_6 + 5H] + 3P(s, \mathrm{white}) \tag{B.7}$$

$$3P(s, \mathrm{white}) + 9(H^+ + e^-) \rightleftharpoons 3PH_3(aq) \tag{B.8}$$

where the first step is the adsorption of H and the desorption of P(s, white) and the second step is the reduction of P(s, white) to $PH_3(aq)$. The DFT total energy change

for the first step is

$$\Delta E^{\mathrm{DFT}}_{\mathrm{dsrp}} = E^{\mathrm{DFT}}_{\mathrm{Ni_9P_6+5H}} + 3E^{\mathrm{DFT}}_{\mathrm{P}(s,\mathrm{white})} - E^{\mathrm{DFT}}_{\mathrm{Ni_9P_6+3P}} - (5/2)E^{\mathrm{DFT}}_{\mathrm{H_2}(g)} = 1.72 \text{ eV} \tag{B.9}$$

The zero point energy change is

$$\Delta\mathrm{ZPE} = \mathrm{ZPE}_{\mathrm{Ni_9P_6+5H}} - \mathrm{ZPE}_{\mathrm{Ni_9P_6+3P}} - (5/2)\mathrm{ZPE}_{\mathrm{H_2}(g)} = 0.26 \text{ eV} \tag{B.10}$$

where $\mathrm{ZPE}_{\mathrm{Ni_9P_6+5H}}$ includes the H adsorbates and the surface atoms coupled to them and $\mathrm{ZPE}_{\mathrm{Ni_9P_6+3P}}$ just includes the same surface atoms. The integrated heat capacity change is

$$\Delta(H - H^\circ) = -(5/2)(H - H^\circ)_{\mathrm{H_2}(g)} = -0.22 \text{ eV} \tag{B.11}$$

The standard molar entropy change is

$$-T\Delta S^\circ = (5/2)TS^\circ_{\mathrm{H_2}(g)} = 1.01 \text{ eV} \tag{B.12}$$

From Eq. B.2, $\Delta G_{\mathrm{dsrp}} = 2.77$ eV. Inserting this value into Eq. 4.10 and noting that $n_{\mathrm{Ni}} = 0$ and $n_{\mathrm{P}} = 3$ (the numbers of Ni and P removed relative to the reference) for this reaction, we arrive at the following equation for the surface free energy relative to the reference surface

$$\begin{aligned} \Delta G_{\mathrm{diss}} &= 2.77 + 3\Delta G^\circ_{\mathrm{P}(s,\mathrm{white})/\mathrm{PH_3}(aq)} + 0.83p\mathrm{H} + 14U \\ &= 2.82 + 0.83p\mathrm{H} + 14U \end{aligned} \tag{B.13}$$

Table B.2: Free energy of $Ni_2P(0001)$ surfaces in equilibrium with 1 M Ni^{2+} or $Ni(s)$, and 1 M PH_3 or 1 M H_3PO_4 at 298.15 K. n corresponds to the number of atoms removed from the reference surface, Ni_2P-Ni_3P_2+P, to obtain a given surface. All energies are reported in eV.

Surface	n_{Ni}	n_P	n_H	ΔE_{dsrp}^{DFT}	ΔZPE	$\Delta(H-H^\circ)$	$-T\Delta S^\circ$	ΔG_{dsrp}
Ni_3P_2+P (reference)	0	0	0	0.00	-	-	-	0.00
Ni_3P_2+P+0.33V_{Ni}	1	0	0	1.37	-	-	-	1.37
Ni_3P_2+0.67P	0	1	0	1.24	-	-	-	1.24
Ni_3P_2+0.33P	0	2	0	2.48	-	-	-	2.48
Ni_3P_2+P+0.33Ni	-1	0	0	-0.39	-	-	-	-0.39
Ni_3P_2+1.33P	0	-1	0	-0.48	-	-	-	-0.48
Ni_3P_2+P+0.33H	0	0	-1	-0.22	0.08	-0.04	0.20	0.01
Ni_3P_2+P+0.67H	0	0	-2	-0.43	0.14	-0.09	0.40	0.03
Ni_3P_2+P+H	0	0	-3	-0.62	0.21	-0.13	0.61	0.07
Ni_3P_2+P+1.33H	0	0	-4	-0.84	0.30	-0.18	0.81	0.10
Ni_3P_2+P+1.67H	0	0	-5	-1.03	0.39	-0.22	1.01	0.16
Ni_3P_2+P+2H	0	0	-6	-1.25	0.48	-0.26	1.21	0.18
Ni_3P_2+P+2.33H	0	0	-7	-1.36	0.57	-0.31	1.41	0.32
Ni_3P_2+P+0.33O	0	0	2	-0.31	-0.19	-0.01	0.18	-0.33
Ni_3P_2+P+0.67O	0	0	4	-0.58	-0.38	-0.03	0.36	-0.63
Ni_3P_2+P+O	0	0	6	-0.81	-0.57	-0.04	0.54	-0.88
Ni_3P_2+P+0.33OH	0	0	1	-0.35	-0.06	-0.06	0.38	-0.08
Ni_3P_2+P+0.67OH	0	0	2	-0.70	-0.12	-0.12	0.76	-0.17
Ni_3P_2+P+OH	0	0	3	-1.05	-0.17	-0.18	1.14	-0.25
Ni_3P_2+0.33HO-P=O	0	0	3	-0.28	-0.24	-0.07	0.56	-0.03
Ni_3P_2+0.33$HP(OH)_2$	0	0	1	-0.75	0.03	-0.16	0.97	0.09
Ni_3P_2+0.33HP=O	0	0	1	0.35	-0.12	-0.06	0.38	0.56
Ni_3P_2+0.33HPOH	0	0	0	-0.54	0.06	-0.10	0.58	0.00
Ni_3P_2+0.33(OH+H)	0	0	0	-0.49	0.06	-0.10	0.58	0.05
Ni_3P_2	0	3	0	3.72	-	-	-	3.72
Ni_3P_2+0.33V_{Ni}	1	3	0	4.64	-	-	-	4.64
Ni_3P_2+0.33V_P	0	4	0	5.81	-	-	-	5.81
Ni_3P_2+0.33Ni	-1	3	0	4.01	-	-	-	4.01
Ni_3P_2+0.33H	0	3	-1	3.05	0.03	-0.04	0.20	3.24
Ni_3P_2+0.67H	0	3	-2	2.40	0.07	-0.09	0.40	2.79
Ni_3P_2+H	0	3	-3	1.75	0.10	-0.13	0.61	2.32
Ni_3P_2+1.33H^α	0	3	-4	1.78	0.18	-0.18	0.81	2.58
Ni_3P_2+1.33H^β	0	3	-4	1.79	0.17	-0.18	0.81	2.59
Ni_3P_2+1.67H^α	0	3	-5	1.97	0.23	-0.22	1.01	3.00
Ni_3P_2+1.67H^β	0	3	-5	1.91	0.20	-0.22	1.01	2.89
Ni_3P_2+1.67H^γ	0	3	-5	1.72	0.26	-0.22	1.01	2.77
Ni_3P_2+1.67H^δ	0	3	-5	1.88	0.24	-0.22	1.01	2.90
Ni_3P_2+0.33O	0	3	2	4.21	-0.21	-0.01	0.18	4.17
Ni_3P_2+0.33OH	0	3	1	3.17	-0.04	-0.06	0.38	3.46
Ni_3P_2+0.67OH	0	3	2	2.63	-0.08	-0.12	0.76	3.19
Ni_3P_2+OH	0	3	3	2.08	-0.12	-0.18	1.14	2.93

Table B.3: Free energy of $Ni_5P_4(0001)$ surfaces relative to Ni_5P_4-$Ni_4P_3(0001)$ in equilibrium with 1 M Ni^{2+} or Ni(s), and 1 M PH_3 or 1 M H_3PO_4 at 298.15 K. All energies are reported in eV.

Surface	n_{Ni}	n_P	n_H	ΔE_{dsrp}^{DFT}	ΔZPE	$\Delta(H-H^\circ)$	$-T\Delta S^\circ$	ΔG_{dsrp}
Ni_4P_3+3P	0	-9	0	-9.23	-	-	-	-9.23
Ni_4P_3+3P+0.33V_{Ni}	1	-9	0	-8.29	-	-	-	-8.29
Ni_4P_3+3P+0.33Ni^{α}	-1	-9	0	-9.74	-	-	-	-9.74
Ni_4P_3+3P+0.33Ni^{β}	-1	-9	0	-9.74	-	-	-	-9.74
Ni_4P_3+3P+0.33H	0	-9	-1	-9.43	0.06	-0.04	0.20	-9.21
Ni_4P_3+3P+0.67H	0	-9	-2	-9.60	0.13	-0.09	0.40	-9.15
Ni_4P_3+3P+H	0	-9	-3	-9.74	0.19	-0.13	0.61	-9.07
Ni_4P_3+3P+1.33H	0	-9	-4	-9.85	0.28	-0.18	0.81	-8.94
Ni_4P_3+3P+1.67H	0	-9	-5	-9.97	0.36	-0.22	1.01	-8.82
Ni_4P_3+3P+2H	0	-9	-6	-10.04	0.45	-0.26	1.21	-8.65
Ni_4P_3+3P+2.33H	0	-9	-7	-10.14	0.53	-0.31	1.41	-8.51
Ni_4P_3+3P+2.67H	0	-9	-8	-10.37	0.62	-0.35	1.62	-8.49
Ni_4P_3+3P+3H	0	-9	-9	-10.44	0.72	-0.39	1.82	-8.30
Ni_4P_3+3P+3.33H	0	-9	-10	-10.61	0.78	-0.44	2.02	-8.25
Ni_4P_3+2.67P	0	-8	0	-8.63	-	-	-	-8.63
Ni_4P_3+2.67P+0.33V_{Ni}	1	-8	0	-7.72	-	-	-	-7.72
Ni_4P_3+2.33P	0	-7	0	-8.00	-	-	-	-8.00
Ni_4P_3+1.67P	0	-5	0	-5.52	-	-	-	-5.52
Ni_4P_3+1.33P	0	-4	0	-3.72	-	-	-	-3.72
Ni_4P_3+P	0	-3	0	-1.92	-	-	-	-1.92
Ni_4P_3+0.67P	0	-2	0	-1.28	-	-	-	-1.28
Ni_4P_3+0.33P	0	-1	0	-0.64	-	-	-	-0.64
Ni_4P_3+2.67P+0.33H	0	-8	-1	-8.83	0.05	-0.04	0.20	-8.63
Ni_4P_3+2.67P+0.67H	0	-8	-2	-9.33	0.13	-0.09	0.40	-8.88
Ni_4P_3+2.67P+H	0	-8	-3	-9.70	0.24	-0.13	0.61	-8.98
Ni_4P_3+2.67P+1.33H	0	-8	-4	-9.86	0.30	-0.18	0.81	-8.93
Ni_4P_3+2.67P+1.67H	0	-8	-5	-10.03	0.38	-0.22	1.01	-8.86
Ni_4P_3+2.67P+2H	0	-8	-6	-10.13	0.46	-0.26	1.21	-8.72
Ni_4P_3+2.67P+2.33H	0	-8	-7	-10.44	0.57	-0.31	1.41	-8.76
Ni_4P_3+2.67P+2.67H	0	-8	-8	-10.30	0.64	-0.35	1.62	-8.39
Ni_4P_3+2.67P+3H	0	-8	-9	-10.69	0.76	-0.39	1.82	-8.51
Ni_4P_3+2.67P+3.33H	0	-8	-10	-10.53	0.84	-0.44	2.02	-8.11
Ni_4P_3+2.67P+0.33Ni	-1	-8	0	-9.04	-	-	-	-9.04
Ni_4P_3+2.67P+0.67Ni	-2	-8	0	-8.61	-	-	-	-8.61
Ni_4P_3+2.67P+0.33Ni+0.33H	-1	-8	-1	-9.39	0.06	-0.04	0.20	-9.17
Ni_4P_3+2.67P+0.33Ni+0.67H	-1	-8	-2	-9.71	0.15	-0.09	0.40	-9.25
Ni_4P_3+2.67P+0.33Ni+H	-1	-8	-3	-9.84	0.22	-0.13	0.61	-9.15
Ni_4P_3+2.67P+0.33Ni+1.33H	-1	-8	-4	-9.99	0.28	-0.18	0.81	-9.08
Ni_4P_3+2.67P+0.33Ni+1.67H	-1	-8	-5	-10.12	0.35	-0.22	1.01	-8.97
Ni_4P_3+2.67P+0.33Ni+2H	-1	-8	-6	-10.16	0.42	-0.26	1.21	-8.80
Ni_4P_3+2.67P+0.33Ni+2.33H	-1	-8	-7	-10.42	0.53	-0.31	1.41	-8.79
Ni_4P_3+2.67P+0.33Ni+2.67H	-1	-8	-8	-10.69	0.62	-0.35	1.62	-8.80
Ni_4P_3+2.67P+0.33Ni+3H	-1	-8	-9	-10.73	0.76	-0.39	1.82	-8.54
Ni_4P_3+2.67P+0.33Ni+3.33H	-1	-8	-10	-10.33	0.85	-0.44	2.02	-7.90
Ni_4P_3+2P	0	-6	0	-7.32	-	-	-	-7.32
Ni_4P_3+2P+0.33Ni	-1	-6	0	-7.12	-	-	-	-7.12
Ni_4P_3+2P+0.33H	0	-6	-1	-7.53	0.03	-0.04	0.20	-7.34
Ni_4P_3+2P+0.67H	0	-6	-2	-7.74	0.05	-0.09	0.40	-7.36
Ni_4P_3+2P+H	0	-6	-3	-7.92	0.09	-0.13	0.61	-7.37
Ni_4P_3+2P+1.33H	0	-6	-4	-8.43	0.18	-0.18	0.81	-7.61
Ni_4P_3+2P+1.67H	0	-6	-5	-8.89	0.28	-0.22	1.01	-7.82
Ni_4P_3+2P+2H	0	-6	-6	-9.36	0.38	-0.26	1.21	-8.03
Ni_4P_3+2P+2.33H	0	-6	-7	-9.73	0.44	-0.31	1.41	-8.19
Ni_4P_3+2P+2.67H	0	-6	-8	-10.09	0.49	-0.35	1.62	-8.34
Ni_4P_3+2P+3H	0	-6	-9	-10.44	0.55	-0.39	1.82	-8.46
Ni_4P_3+2P+3.33H	0	-6	-10	-10.08	0.57	-0.44	2.02	-7.93
Ni_4P_3 (reference)	0	0	0	0.00	-	-	-	0.00

Ni_4P_3+0.33V_{Ni}	1	0	0	0.61	-	-	-	0.61
Ni_4P_3+0.33V_P	0	1	0	2.36	-	-	-	2.36
Ni_4P_3+0.33Ni^{α}	-1	0	0	-0.49	-	-	-	-0.49
Ni_4P_3+0.33Ni^{β}	-1	0	0	-0.48	-	-	-	-0.48
Ni_4P_3+0.67Ni	-2	0	0	-0.96	-	-	-	-0.96
Ni_4P_3+0.33H	0	0	-1	-0.55	0.05	-0.04	0.20	-0.34
Ni_4P_3+0.67H	0	0	-2	-1.07	0.11	-0.09	0.40	-0.64
Ni_4P_3+H	0	0	-3	-1.58	0.14	-0.13	0.61	-0.97
Ni_4P_3+1.33H	0	0	-4	-2.19	0.23	-0.18	0.81	-1.34
Ni_4P_3+1.67H	0	0	-5	-2.77	0.31	-0.22	1.01	-1.67
Ni_4P_3+2H	0	0	-6	-3.33	0.40	-0.26	1.21	-1.99
Ni_4P_3+2.33H	0	0	-7	-3.86	0.48	-0.31	1.41	-2.27
Ni_4P_3+2.67H	0	0	-8	-4.35	0.57	-0.35	1.62	-2.52
Ni_4P_3+3H	0	0	-9	-4.87	0.66	-0.39	1.82	-2.79
Ni_4P_3+3.33H	0	0	-10	-5.35	0.76	-0.44	2.02	-3.01
Ni_4P_3+3.67H	0	0	-11	-5.87	0.73	-0.48	2.22	-3.40
Ni_4P_3+4H	0	0	-12	-6.36	0.84	-0.53	2.42	-3.63
Ni_4P_3+4.33H	0	0	-13	-6.26	0.88	-0.57	2.62	-3.32
Ni_4P_3+Ni	-3	0	0	-1.37	-	-	-	-1.37
Ni_4P_3+1.33Ni	-4	0	0	-0.42	-	-	-	-0.42
Ni_4P_3+Ni+0.33H	-3	0	-1	-1.99	0.05	-0.04	0.20	-1.79
Ni_4P_3+Ni+0.67H	-3	0	-2	-2.55	0.10	-0.09	0.40	-2.14
Ni_4P_3+Ni+H	-3	0	-3	-3.14	0.14	-0.13	0.61	-2.52
Ni_4P_3+Ni+1.33H	-3	0	-4	-2.99	0.18	-0.18	0.81	-2.18
Ni_4P_3+Ni+1.67H	-3	0	-5	-3.44	0.25	-0.22	1.01	-2.40
Ni_3P_3	21	15	0	21.77	-	-	-	21.77
Ni_3P_3+0.33V_{Ni}	22	15	0	22.79	-	-	-	22.79
Ni_3P_3+0.33V_P	21	16	0	22.58	-	-	-	22.58
Ni_3P_3+0.33Ni	20	15	0	21.95	-	-	-	21.95
Ni_3P_3+0.33P	21	14	0	20.92	-	-	-	20.92
Ni_3P_3+0.33H	21	15	-1	21.04	0.06	-0.04	0.20	21.25
Ni_3P_3+0.67H	21	15	-2	20.31	0.12	-0.09	0.40	20.74
Ni_3P_3+H	21	15	-3	19.64	0.18	-0.13	0.61	20.29
Ni_3P_3+1.33H	21	15	-4	19.63	0.26	-0.18	0.81	20.52
Ni_3P_3+1.67H	21	15	-5	19.65	0.34	-0.22	1.01	20.78
Ni_3P_3'	21	15	0	19.96	-	-	-	19.96
Ni_3P_3'+0.33V_{Ni}	22	15	0	20.78	-	-	-	20.78
Ni_3P_3'+0.33Ni	20	15	0	20.14	-	-	-	20.14
Ni_3P_3'+0.33H	21	15	-1	19.71	0.07	-0.04	0.20	19.94
Ni_3P_3'+0.67H	21	15	-2	19.50	0.15	-0.09	0.40	19.97
Ni_3P_3'+H	21	15	-3	19.33	0.22	-0.13	0.61	20.03
Ni_3P_3'+1.33H	21	15	-4	19.30	0.28	-0.18	0.81	20.21
Ni_3P_3'+1.67H	21	15	-5	19.01	0.33	-0.22	1.01	20.13
Ni_3P_3'+2H	21	15	-6	19.05	0.44	-0.26	1.21	20.44
Ni_3P_3'+V_P	21	18	0	21.31	-	-	-	21.31
Ni_3P_3'+V_P+0.33V_{Ni}	22	18	0	22.57	-	-	-	22.57
Ni_3P_3'+1.33V_P	21	19	0	22.84	-	-	-	22.84
Ni_3P_3'+1.67V_P	21	20	0	24.36	-	-	-	24.36
Ni_3P_3'+V_P+0.33Ni	20	18	0	22.01	-	-	-	22.01
Ni_3P_3'+V_P+0.33H	21	18	-1	21.11	0.05	-0.04	0.20	21.32
Ni_3P_3'+V_P+0.67H	21	18	-2	20.94	0.10	-0.09	0.40	21.36
Ni_3P_3'+V_P+H	21	18	-3	20.78	0.15	-0.13	0.61	21.41
Ni_3P_3'+V_P+1.33H	21	18	-4	20.67	0.21	-0.18	0.81	21.51
Ni_3P_3'+V_P+1.67H	21	18	-5	20.55	0.27	-0.22	1.01	21.62
Ni_3P_3'+V_P+2H	21	18	-6	20.43	0.34	-0.26	1.21	21.72
Ni_3P_3'+V_P+2.33H	21	18	-7	20.42	0.42	-0.31	1.41	21.95
Ni_3P_3'+V_P+2.67H	21	18	-8	20.46	0.48	-0.35	1.62	22.20
Ni_3P_3'+2V_P	21	21	0	25.87	-	-	-	25.87
Ni_3P_3'+2.33V_P	21	22	0	28.58	-	-	-	28.58
Ni_3P_3'+2V_P+0.33Ni	20	21	0	26.47	-	-	-	26.47
Ni_3P_3'+2V_P+0.33H	21	21	-1	25.10	0.04	-0.04	0.20	25.29
Ni_3P_3'+2V_P+0.67H	21	21	-2	24.33	0.09	-0.09	0.40	24.73
Ni_3P_3'+2V_P+H	21	21	-3	23.55	0.14	-0.13	0.61	24.16
Ni_3P_3'+2V_P+1.33H	21	21	-4	23.47	0.18	-0.18	0.81	24.28
Ni_3P_3'+2V_P+1.67H	21	21	-5	23.40	0.22	-0.22	1.01	24.41
Ni_3P_3'+2V_P+2H	21	21	-6	23.33	0.26	-0.26	1.21	24.54

$Ni_3P'_3$+2V_P+2.33H	21	21	-7	23.30	0.33	-0.31	1.41	24.73
$Ni_3P'_3$+2V_P+2.67H	21	21	-8	23.27	0.39	-0.35	1.62	24.92
$Ni_3P'_3$+2V_P+3H	21	21	-9	23.25	0.45	-0.39	1.82	25.12
$Ni_3P'_3$+2V_P+3.33H	21	21	-10	23.36	0.52	-0.44	2.02	25.46

Table B.4: Free energy of $Ni_5P_4(000\bar{1})$ surfaces relative to Ni_5P_4-$Ni_4P_3(000\bar{1})$ in equilibrium with 1 M Ni^{2+} or Ni(s), and 1 M PH_3 or 1 M H_3PO_4 at 298.15 K. All energies are reported in eV.

Surface	n_{Ni}	n_P	n_H	ΔE_{dsrp}^{DFT}	ΔZPE	$\Delta(H-H^\circ)$	$-T\Delta S^\circ$	ΔG_{dsrp}
Ni_4P_3+3P	0	-9	0	-10.90	-	-	-	-10.90
Ni_4P_3+3P+0.33V_{Ni}	1	-9	0	-9.87	-	-	-	-9.87
Ni_4P_3+2.67P	0	-8	0	-10.16	-	-	-	-10.16
Ni_4P_3+2.33P	0	-7	0	-9.45	-	-	-	-9.45
Ni_4P_3+2P	0	-6	0	-8.79	-	-	-	-8.79
Ni_4P_3+1.67P	0	-5	0	-6.72	-	-	-	-6.72
Ni_4P_3+1.33P	0	-4	0	-4.65	-	-	-	-4.65
Ni_4P_3+P	0	-3	0	-2.57	-	-	-	-2.57
Ni_4P_3+0.67P	0	-2	0	-1.72	-	-	-	-1.72
Ni_4P_3+0.33P	0	-1	0	-0.86	-	-	-	-0.86
Ni_4P_3+3P+0.33Ni	-1	-9	0	-11.15	-	-	-	-11.15
Ni_4P_3+3P+0.67Ni	-2	-9	0	-11.01	-	-	-	-11.01
Ni_4P_3+2P+0.33Ni	-1	-6	0	-8.50	-	-	-	-8.50
Ni_4P_3+3P+0.33H	0	-9	-1	-11.17	0.08	-0.04	0.20	-10.93
Ni_4P_3+3P+0.67H	0	-9	-2	-11.40	0.16	-0.09	0.40	-10.93
Ni_4P_3+3P+H	0	-9	-3	-11.59	0.24	-0.13	0.61	-10.87
Ni_4P_3+3P+1.33H	0	-9	-4	-11.77	0.33	-0.18	0.81	-10.81
Ni_4P_3+3P+1.67H	0	-9	-5	-11.88	0.38	-0.22	1.01	-10.71
Ni_4P_3+3P+2H	0	-9	-6	-11.98	0.45	-0.26	1.21	-10.58
Ni_4P_3+3P+2.33H	0	-9	-7	-12.04	0.57	-0.31	1.41	-10.36
Ni_4P_3+3P+2.67H	0	-9	-8	-12.10	0.68	-0.35	1.62	-10.15
Ni_4P_3+3P+3H	0	-9	-9	-12.13	0.78	-0.39	1.82	-9.93
Ni_4P_3 (reference)	0	0	0	0.00	-	-	-	0.00
Ni_4P_3+0.33V_{Ni}	1	0	0	0.67	-	-	-	0.67
Ni_4P_3+0.33V_P	0	1	0	2.96	-	-	-	2.96
Ni_4P_3+0.33Ni	-1	0	0	-0.44	-	-	-	-0.44
Ni_4P_3+0.33H	0	0	-1	-0.78	0.05	-0.04	0.20	-0.57
Ni_4P_3+0.67H	0	0	-2	-1.57	0.10	-0.09	0.40	-1.16
Ni_4P_3+H	0	0	-3	-2.34	0.14	-0.13	0.61	-1.72
Ni_4P_3+1.33H	0	0	-4	-2.88	0.25	-0.18	0.81	-1.99
Ni_4P_3+1.67H	0	0	-5	-3.31	0.35	-0.22	1.01	-2.16
Ni_4P_3+2H	0	0	-6	-3.76	0.46	-0.26	1.21	-2.36
Ni_4P_3+2.33H	0	0	-7	-4.20	0.55	-0.31	1.41	-2.55
Ni_4P_3+2.67H	0	0	-8	-4.51	0.64	-0.35	1.62	-2.61
Ni_4P_3+3H	0	0	-9	-4.89	0.74	-0.39	1.82	-2.73
Ni_4P_3+3.33H	0	0	-10	-5.28	0.85	-0.44	2.02	-2.85
Ni_4P_3+3.67H	0	0	-11	-5.63	0.96	-0.48	2.22	-2.93
Ni_4P_3+4H	0	0	-12	-5.96	1.07	-0.53	2.42	-3.00
Ni_4P_3+4.33H	0	0	-13	-5.75	1.10	-0.57	2.62	-2.59
Ni_4P_3+4.67H	0	0	-14	-5.69	1.22	-0.61	2.83	-2.26

B.5. Structure and Aqueous Stability of $Ni_5P_4(0001)$ Surfaces

For $Ni_5P_4(0001)$, five different surface phases are observed under acidic (pH = -0.1–1), reducing (U = 0.0 to -0.8 V) conditions (see Fig. B.3A). For U > -0.12 V, the surface layer has a stoichiometry of $Ni_5P_4(s)/Ni_3P_3(0001)+2V_P+H$. A representative structure for the family of $Ni_5P_4(s)/Ni_3P_3(0001)+2V_P+nH$ surfaces with n = 1, 2, and 10/3 is shown in Fig. B.3B. This surface comprises Ni_3-hollows that are connected via P atoms, generating a honeycomb-like surface lattice. Similar to the other surface terminations discussed, an H atom binds strongly to each Ni_3-hollow (labelled 1-3), with a bond length of 1.74 Å and ΔG_{ads} = -0.57 eV/H. For -0.12 V $\geq U \geq$ -0.16 V, H atoms populate adjacent Ni-P bridge sites (labelled 4-6) forming $Ni_5P_4(s)/Ni_3P_3(0001)+2V_P+2H$.

For -0.16 V $\geq U \geq$ -0.3 V, the predominant surface phase is $Ni_5P_4(s)/Ni_4P_3(0001)+4H$. This dramatic change in surface composition is driven by surface dissolution, which forms Ni^{2+} and H_3PO_4. The structure of $Ni_5P_4(s)/Ni_4P_3(0001)+nH$ is shown in Fig. B.3C which exhibits repeating Ni_3- and P_3-hollows with central Ni atoms. There are two different types of Ni-P bonds, one with a bond length of 2.11 Å between P and the central Ni and another with a bond length of 2.27 Å between P and a Ni from a Ni_3-hollow. Three H atoms adsorb per $\sqrt{3} \times \sqrt{3}$ supercell, one at each Ni_3-hollow, shown by the inset. The hydrogen adsorbate forms bonds of equal length with all three Ni atoms (1.73 Å). This is shorter than the Ni-H bond length at Ni_3-hollows on $Ni_2P(s)/Ni_3P_2(0001)+H$, 1.79 Å (see Fig. 4.1B). H binding is weaker at the Ni_3-hollow site on $Ni_5P_4(s)/Ni_4P_3(0001)+4H$, with an average adsorption free energy of -0.32 eV/H. Nine additional H atoms per supercell adsorb, three per P_3-hollow site. Each H makes a single P-H bond of length 1.42 Å, and the H atoms point toward the center of the P_3-hollow. For -0.30 V $\geq U \geq$

-0.48 V, one additional H atom per supercell binds to one of the Ni_3-hollows (n = 13/3), resulting in two H atoms binding at the Ni-Ni bridge sites, as shown by the inset in Fig. B.3C.

$U \leq$ -0.48 V reforms $Ni_5P_4(s)/Ni_3P_3(0001)+2V_P$ saturated with H, +(10/3)H. The structure of this phase can be found in Fig. B.3B. In addition to the H atoms at Ni_3-hollow (labelled 1-3) and Ni-P bridge sites (4-6), four more H atoms (7-10) bind weakly to lattice P (ΔG_{ads} ranging from 0.12 to 0.34 eV) at Ni-P bridge sites ($\bar{d}_{\mathrm{PH}}$ = 1.54 Å).

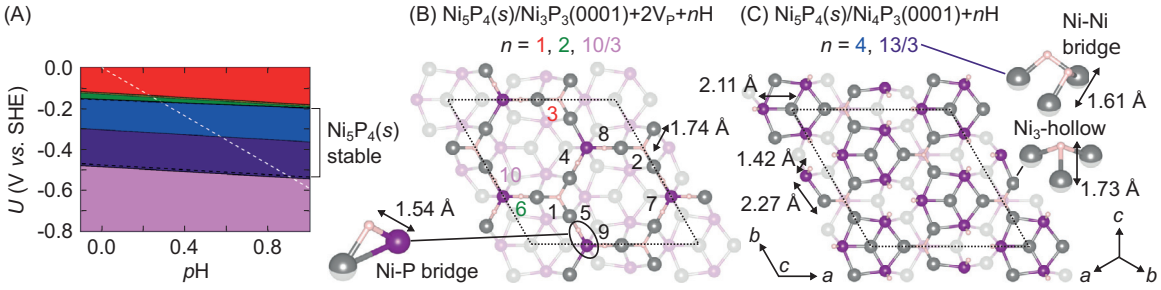

Figure B.3: (A) Surface phase diagram of $Ni_5P_4(0001)$ in equilibrium with 1 M Ni^{2+} or $Ni(s)$, and 1 M PH_3 or 1 M H_3PO_4 at 298.15 K. Evolution of the (B) $Ni_5P_4(s)/Ni_3P_3(0001)+2V_P+nH$ and (C) $Ni_5P_4(s)/Ni_4P_3(0001)+nH$ surfaces through adsorption-desorption equilibrium of H, P, and Ni. n corresponds to the number of H atoms per 1×1 surface unit cell. Numbers next to H atoms correspond to the number of H atoms per $\sqrt{3} \times \sqrt{3}$ $R30°$ surface unit cell. Red, green, and pink regions: $Ni_5P_4(s)/Ni_3P_3(0001)+2V_P+nH$ with $n = 1$, 2, and 10/3, respectively. Blue and violet regions: $Ni_5P_4(s)/Ni_4P_3(0001)+nH$ with $n = 4$ and 13/3, respectively. Some average bond lengths are reported.

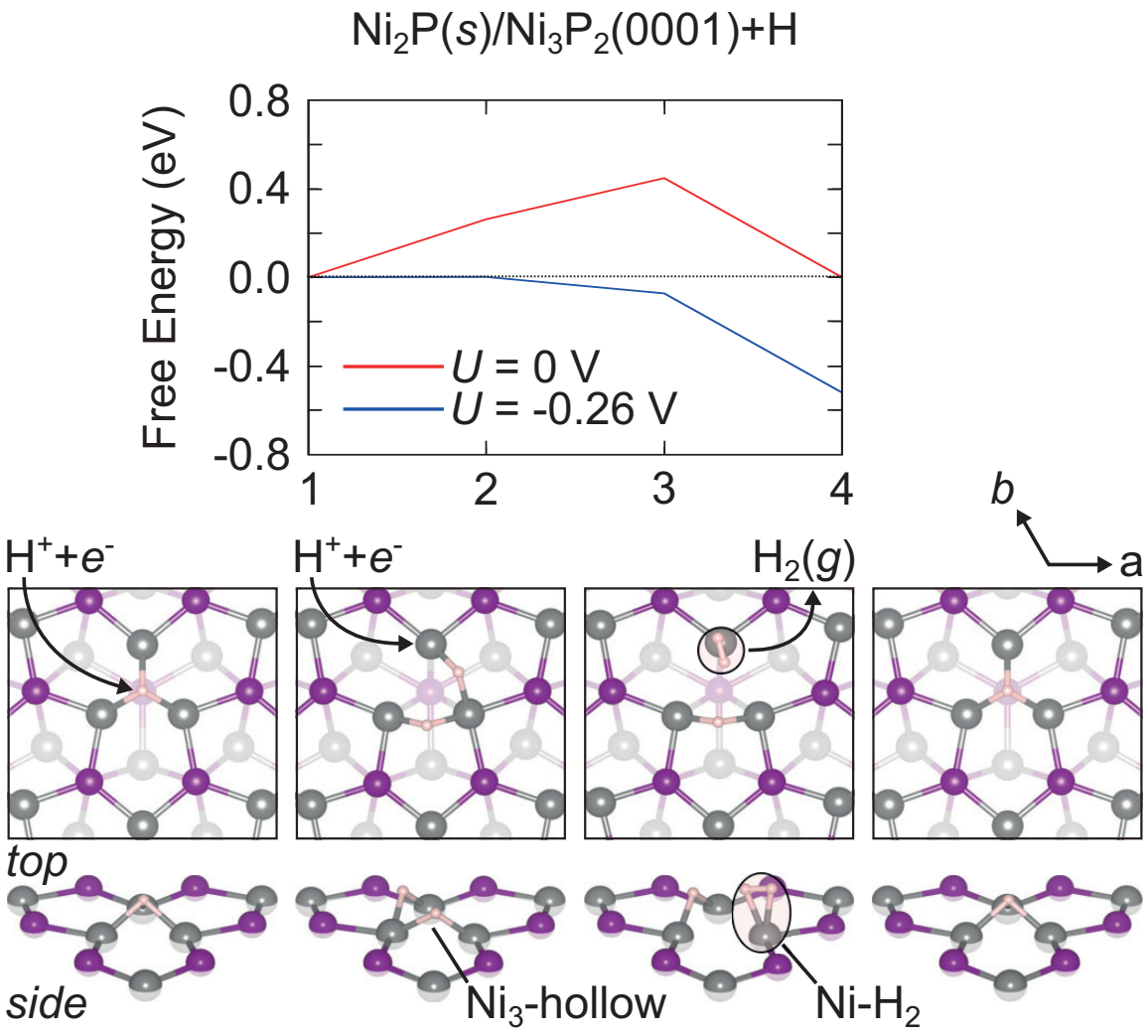

Figure B.4: Free energies and structures of intermediates in the HER for $Ni_2P(s)/Ni_3P_2(0001)$+H in equilibrium with 1 M Ni^{2+} or $Ni(s)$, and 1 M PH_3 or 1 M H_3PO_4 at 298.15 K and pH = 0. The blue line corresponds to the minimum overpotential to make the reaction spontaneous and ensure catalyst stability.

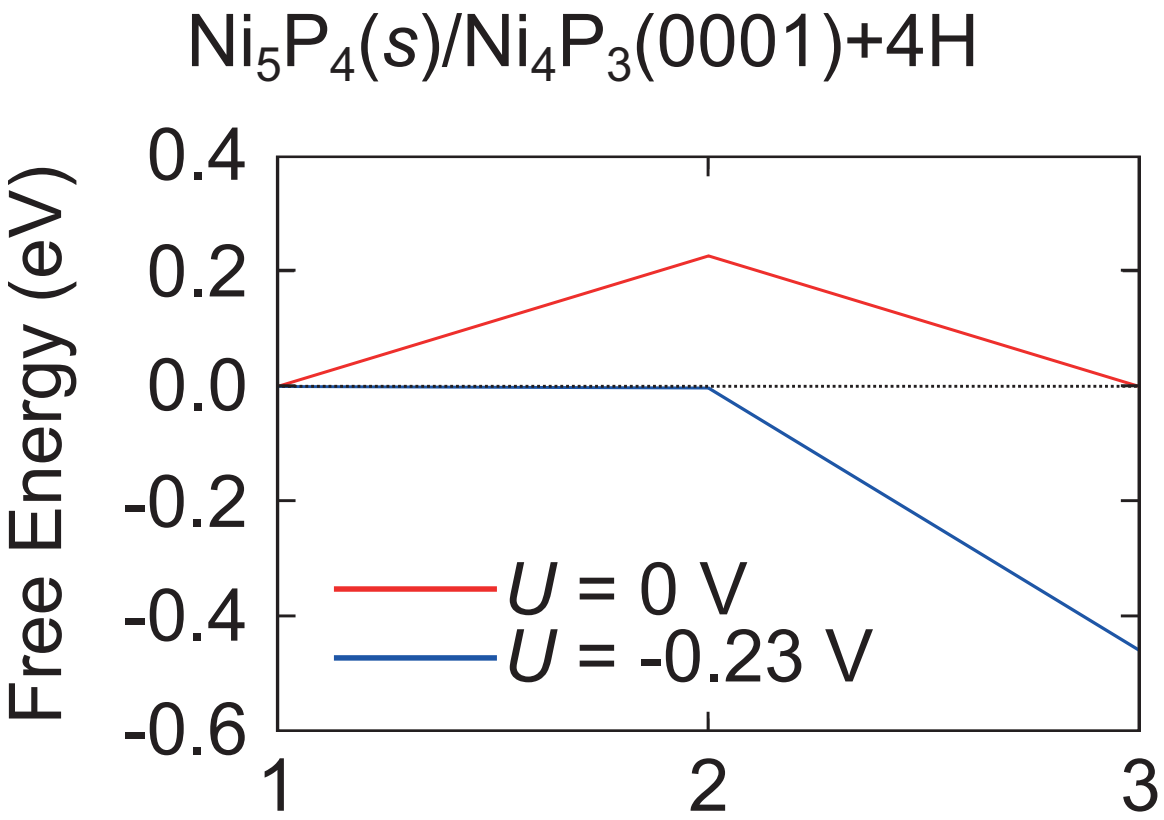

Figure B.5: Free energy of intermediates in the HER for $Ni_5P_4(0001)$ in equilibrium with 1 M Ni^{2+} or $Ni(s)$, and 1 M PH_3 or 1 M H_3PO_4 at 298.15 K and pH = 0. The blue line corresponds to the minimum overpotential to make the reaction spontaneous and ensure catalyst stability. The structure of intermediates in the HER is identical to that of $Ni_5P_4(000\bar{1})$ as shown in Fig. 4.4B.

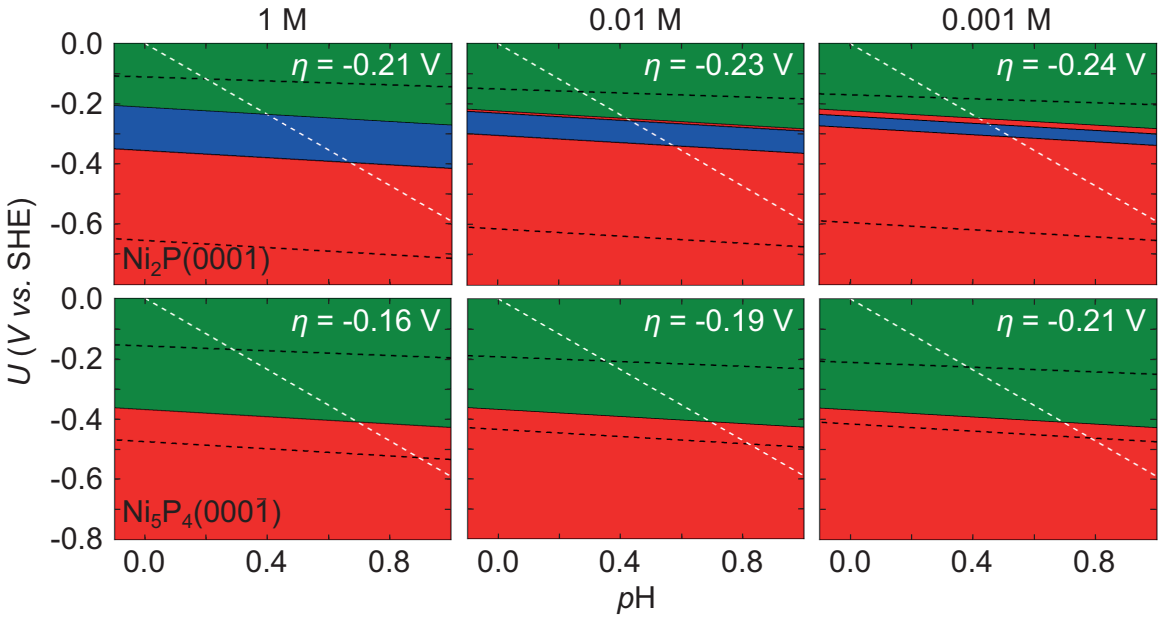

Figure B.6: Surface phase diagram of $Ni_2P(0001)$ and $Ni_5P_4(000\bar{1})$ in equilibrium with different molar concentrations of solvated species at 298.15 K and pH = 0. Black dashed lines enclose the bulk stability regions for Ni_2P and Ni_5P_4. At conditions (U and pH) below the white dashed line, water is unstable and $H_2(g)$ evolution is thermodynamically favored. For $Ni_2P(0001)$, both the bulk and $Ni_2P(s)/Ni_3P_2(0001)$+P+(7/3)H surface (blue region) stability domains shrink with respect to the applied voltage. For $Ni_5P_4(000\bar{1})$, only the bulk stability region shrinks with respect to the applied voltage. η corresponds to the minimum overpotential to make the reaction spontaneous and ensure catalyst stability.

APPENDIX C : Supplemental: Tuning the H_2 evolving activity of Ni_2P via surface nonmetal doping-generated chemical pressure: a joint first principles and machine learning study

C.1. Additional Computational Details

C.1.1. DFT

We applied Gaussian electronic smearing of 0.07 eV to the band occupations near the Fermi energy to improve electronic k-point convergence. The total energy convergence threshold for SCF calculations was 1.4×10^{-5} eV/cell. The total energy and force convergence thresholds for bulk and surface geometry relaxations were 1.4×10^{-3} eV/cell and 2.6×10^{-2} eV/Å, respectively. Pseudopotentials were constructed using the OPIUM (version 3.7) software.

C.1.2. Machine Learning

Atomic radii were taken from Ref. 340. The dimensionality of the data set for dopant residual charge descriptors is six (corresponding to the maximum substitution number, $n_X = 6$). For $n_X < 6$, there will naturally be less than six values for the dopant residual charge. In the aforementioned cases, we replace the missing data with "-999", which the algorithm will interpret as "completely different". We also removed zero and near zero-variance descriptors. Fig. C.1 is a schematic for 3-fold cross-validation (CV). Prior to training, the observations are reordered randomly so that the distribution of dopant identities and concentrations is consistent with that of the full data set. In the first iteration, the first third of the data is selected as the test data and

the machine learning algorithm is trained on the remaining two-thirds of the data. In the second iteration, the second third of the data is selected as the test data and so on.

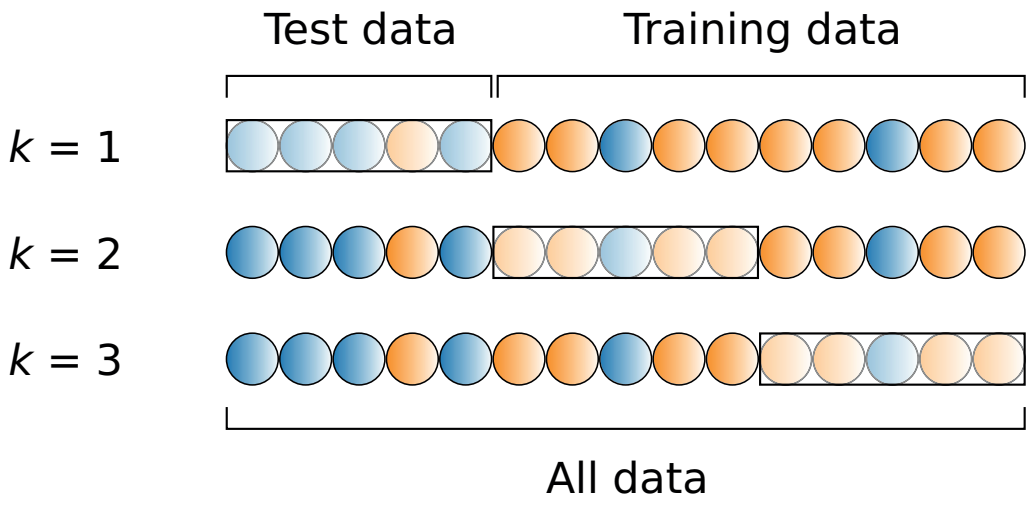

Figure C.1: Schematic for 3-fold cross-validation adapted from Ref. 5. The data set is represented as a chain of circles whose color, blue or orange, corresponds to their class or y-value. For example, blue and orange could represent $\Delta G_{\mathrm{H}} \leq 0$ and $\Delta G_{\mathrm{H}} > 0$, respectively. The test data is enclosed by a rectangle for each k.

C.2. Additional Theoretical Details

C.2.1. Free Energy of H Adsorption

For surface H adsorption, *i.e.*

$$\mathrm{Su} + \mathrm{H}^+ + e^- \rightleftharpoons \mathrm{Su} - \mathrm{H} \tag{C.1}$$

the free energy change (ΔG_H) can be written as

$$\Delta G_\mathrm{H} = G_\mathrm{Su-H} - G_\mathrm{Su} - \frac{1}{2} G_{\mathrm{H}_2} \tag{C.2}$$

where Su is the surface. This can be rewritten in calculable and experimentally-obtainable variables as

$$\begin{aligned} \Delta G_\mathrm{H} = \Delta E_\mathrm{H}^\mathrm{DFT} + \Delta \mathrm{ZPE}_\mathrm{H} - \frac{1}{2}(H - H^\circ)_{\mathrm{H}_2} + \frac{1}{2} T S^\circ_{\mathrm{H}_2} \\ + 2.303 n_\mathrm{H} k_\mathrm{B} T p\mathrm{H} + n_e q_e U \end{aligned} \tag{C.3}$$

where T is the temperature in Kelvin, U is the electrode potential *vs.* SHE in volts, E^DFT is the DFT total energy, ZPE is the zero-point energy, $H - H^\circ$ is the integrated heat capacity, S^0 is the standard entropy, n_H is the number of protons consumed, k_B is Boltzmann's constant, n_e is the number of electrons consumed, and q_e is the electron charge. Here, we assume that surface pV and TS contributions are negligible. Additionally, it has been shown that for metal-hollow sites (38; 341), ΔG_H at 300 K can be approximated as

$$\Delta G_\mathrm{H}(300\ \mathrm{K}) = \Delta E_\mathrm{H}^\mathrm{DFT} + 0.24\ \mathrm{eV} + 2.303 n_\mathrm{H} k_\mathrm{B} T \ln p\mathrm{H} + n_e q_e U \tag{C.4}$$

where

$$\Delta \mathrm{ZPE_H} - \frac{1}{2}(H - H^\circ)_{\mathrm{H_2}} + \frac{1}{2} T S^\circ_{\mathrm{H_2}} = 0.24 \text{ eV} \tag{C.5}$$

The above term is demonstrated to be nearly fixed due to the dominance of the nearly constant M-H stretching and bending frequencies on the surface H atom's ZPE. The sum of the frequencies of the M-H normal modes is ≈ 3550 cm^{-1}, corresponding to $\hbar\omega/2 = 0.22$ eV. Variation in the M-H stretching and bending frequencies even by 200 cm^{-1} results in only ≈ 0.01 eV change in the ZPE. Since $\mathrm{ZPE_{H_2}}/2 = 0.14$ eV and $-(H - H^\circ)_{\mathrm{H_2}}/2 + TS^\circ_{\mathrm{H_2}}/2 = 0.16$ eV, the term above is therefore ≈ 0.24 eV.

C.2.2. Free Energy of Dopant Substitution

The dissolution of a surface atom in aqueous solution can be written as

$$\mathrm{Su - A} + n_w \mathrm{H_2O} \rightleftharpoons \mathrm{Su} + [\mathrm{H}_x\mathrm{AO}_y]^z + n_\mathrm{H}\mathrm{H}^+ + n_e e^- \tag{C.6}$$

where n_w is the number of water molecules needed for redox and/or solvation of A, $[\mathrm{H}_x\mathrm{AO}_y]^z$ is the most stable aqueous phase of A, and z is the charge of this phase. Unlike in Eq. C.4, n_H and n_e are the number of protons and electrons released, respectively. Note that we could rewrite n_w, n_H, and n_e in terms of x, y, and z as $n_w = y$, $n_\mathrm{H} = -x + 2y$, and $n_e = -x + 2y + z$. Previously, we derived an expression for the free energy change of this reaction (38), which is

$$\Delta G_{\mathrm{A,diss}} = \Delta G_{\mathrm{dsrp}} + \Delta G^\circ_{\mathrm{A(std)}/\mathrm{H}_x\mathrm{AO}_y^z} + k_\mathrm{B}T \ln a_{\mathrm{H}_x\mathrm{AO}_y^z} - 2.303 n_\mathrm{H} k_\mathrm{B} T p\mathrm{H} - n_e q_e U \tag{C.7}$$

Here, ΔG_{dsrp} is the differential desorption (dsrp) free energy for A to leave the surface and form A(std), $G^\circ_{\mathrm{A(std)}/\mathrm{H}_x\mathrm{AO}_y^z}$ is the standard free energy for redox and/or solvation of A (see Table C.1), and a is the activity. Conversely, the precipitation of an aqueous

solute can be written as

$$\mathrm{Su} + [\mathrm{H}_{x'}\mathrm{XO}_{y'}]^{z'} + n'_{\mathrm{H}}\mathrm{H}^{+} + n'_{e}e^{-} \rightleftharpoons \mathrm{Su} - \mathrm{X} + n'_{w}\mathrm{H}_2\mathrm{O} \tag{C.8}$$

where the $'$ indicates that the stoichiometry of dissolution and precipitation are, in general, not equal. The substitution of surface atom A with X is simply the sum of Eqs. C.6 and C.8, which is

$$\begin{aligned}\mathrm{Su} - \mathrm{A} + [\mathrm{H}_{x'}\mathrm{XO}_{y'}]^{z'} + (n_w - n'_w)\mathrm{H}_2\mathrm{O} \rightleftharpoons \\ \mathrm{Su} - \mathrm{X} + [\mathrm{H}_x\mathrm{AO}_y]^{z} + (n_{\mathrm{H}} - n'_{\mathrm{H}})\mathrm{H}^{+} + (n_e - n'_e)e^{-}\end{aligned} \tag{C.9}$$

Using Eqs. C.7 and C.9, it is straightforward to show that

$$\Delta G_{\mathrm{sub}} = \Delta G^{\circ}_{\mathrm{sub}} + k_{\mathrm{B}}T \ln \frac{\mathrm{H}_x\mathrm{AO}_y^z}{\mathrm{H}_{x'}\mathrm{XO}_{y'}^{z'}} - 2.303 \cdot (n_{\mathrm{H}} - n'_{\mathrm{H}}) \cdot k_{\mathrm{B}}T p\mathrm{H} - (n_e - n'_e) \cdot U \tag{C.10}$$

where we have defined

$$\Delta G^{\circ}_{\mathrm{sub}} = (\Delta G_{\mathrm{dsrp,A}} - \Delta G_{\mathrm{dsrp,X}}) + (\Delta G^{\circ}_{\mathrm{A(std)}/\mathrm{H}_x\mathrm{AO}_y^z} - \Delta G^{\circ}_{\mathrm{X(std)}/\mathrm{H}_{x'}\mathrm{XO}_{y'}^{z'}}) \tag{C.11}$$

as the standard free energy of substitution relative to the most stable phases of A and X at $U = 0$ V *vs.* SHE and $p\mathrm{H} = 0$ (see Table C.2). For A = P and X = dopant, we select $a_{\mathrm{H_3PO_4}}/a_{\mathrm{H}_{x'}\mathrm{XO}_{y'}^{z'}} = 1 \times 10^3$ so that $\mathrm{Ni_2P}$ and not Ni-X is the most stable bulk phase. Note, however, that we set the activity of solids (As, Te, and $\mathrm{SiO_2}$) and liquids ($\mathrm{H_2O}$) to be 1, in which case $a_{\mathrm{H_3PO_4}}/a_{\mathrm{H}_{x'}\mathrm{XO}_{y'}^{z'}} = 1$.

Table C.1: Standard oxidation/reduction free energy of nonmetal standard states to form most stable phases at $U = 0$ V *vs.* SHE and $p\mathrm{H} = 0$ ($\Delta G^{\circ}_{\mathrm{A(std)/H}_x\mathrm{AO}_y^z}$) (1). All energies are reported in eV/molecule.

X	Phase	$\Delta G^{\circ}_{\mathrm{A(std)/H}_x\mathrm{AO}_y^z}$
S	H_2S	-0.29
Se	H_2Se	0.16
Te	Te	0.00
B	H_3BO_3	-2.67
C	CH_4	-0.36
N	NH_4^+	-0.82
O	H_2O	-2.46
Si	SiO_2	-3.90
As	As	0.00
P	H_3PO_4	-2.01

Table C.2: DFT-calculated free energy of substituting P with other nonmetals relative to their standard states ($\Delta G_{\mathrm{dsrp,A}} - \Delta G_{\mathrm{dsrp,X}}$). All energies are reported in eV/atom. Note that DFT-PBE is known to overbind O_2 by 0.80 eV (2) and N_2 by 0.50 eV (3). We corrected this by performing DFT total energy calculations for triplet O and quadruplet N and then adding their experimental binding energies (5.16 eV/O_2 and 9.81 eV/N_2). For example, for O_2 we used $E_{O_2} = 2E_{O}^{\mathrm{DFT}} + E_{\mathrm{bind}}^{\mathrm{exp}}$. For O_2 and N_2, we also included the experimental integrated heat capacity and the standard entropy of the gas.

n_X	As	Si	B	C	N	O	S	Se	Te
1	0.32	0.81	1.54	2.75	2.29	-0.33	-0.25	-0.23	-0.12
2	0.29	0.85	1.43	2.63	2.13	-0.64	-0.37	-0.34	-0.20
3	0.16	0.78	1.61	2.84	2.02	-0.73	-0.48	-0.39	-0.22
4	0.62	0.89	1.51	3.10	2.76	0.78	0.34	0.61	0.92
5	0.64	0.85	1.64	3.35	3.08	0.68	0.35	0.64	1.03
6	0.61	0.82	1.55	3.58	2.95	0.70	0.34	0.74	1.21

Example: Substituting P with S at $T = 300$ K, $U = 0$ V *vs.* SHE, and *p*H = 0

At $T = 300$ K, $U = 0$ V *vs.* SHE, and $p\mathrm{H} = 0$, the most stable aqueous phases of P and S are H_3PO_4 and H_2S, respectively. Therefore, the substitution of P with S can be written as

$$\mathrm{Su} - \mathrm{P} + \mathrm{H_2S} + 4\mathrm{H_2O} \rightleftharpoons \mathrm{Su} - \mathrm{S} + \mathrm{H_3PO_4} + 7(\mathrm{H^+} + e^-) \tag{C.12}$$

The free energy change for this reaction is

$$\begin{aligned}\Delta G_{\mathrm{sub}} = (\Delta G_{\mathrm{dsrp,A}} - \Delta G_{\mathrm{dsrp,X}}) + (\Delta G^{\circ}_{\mathrm{P}(s,\mathrm{white})/\mathrm{H_3PO_4}} - \Delta G^{\circ}_{\mathrm{S}(s,\alpha)/\mathrm{H_2S}}) \\ + k_{\mathrm{B}}T \ln \frac{a_{\mathrm{H_3PO_4}}}{a_{\mathrm{H_2S}}} \\ - 2.303 \cdot 7 \cdot k_{\mathrm{B}}T p\mathrm{H} - 7U\end{aligned} \tag{C.13}$$

Plugging in the conditions and using Tables C.1 and C.2 for $n_{\mathrm{S}} = 1$, we obtain

$$\Delta G_{\mathrm{sub}} = -0.25 + (-2.01 + 0.29) + 0.026 \cdot \ln \frac{a_{\mathrm{H_3PO_4}}}{a_{\mathrm{H_2S}}} \text{ eV} \tag{C.14}$$

For $a_{\mathrm{H_3PO_4}}/a_{\mathrm{H_2S}} = 1 \times 10^3$,

$$\Delta G_{\mathrm{sub}} = -1.79 \text{ eV} \tag{C.15}$$

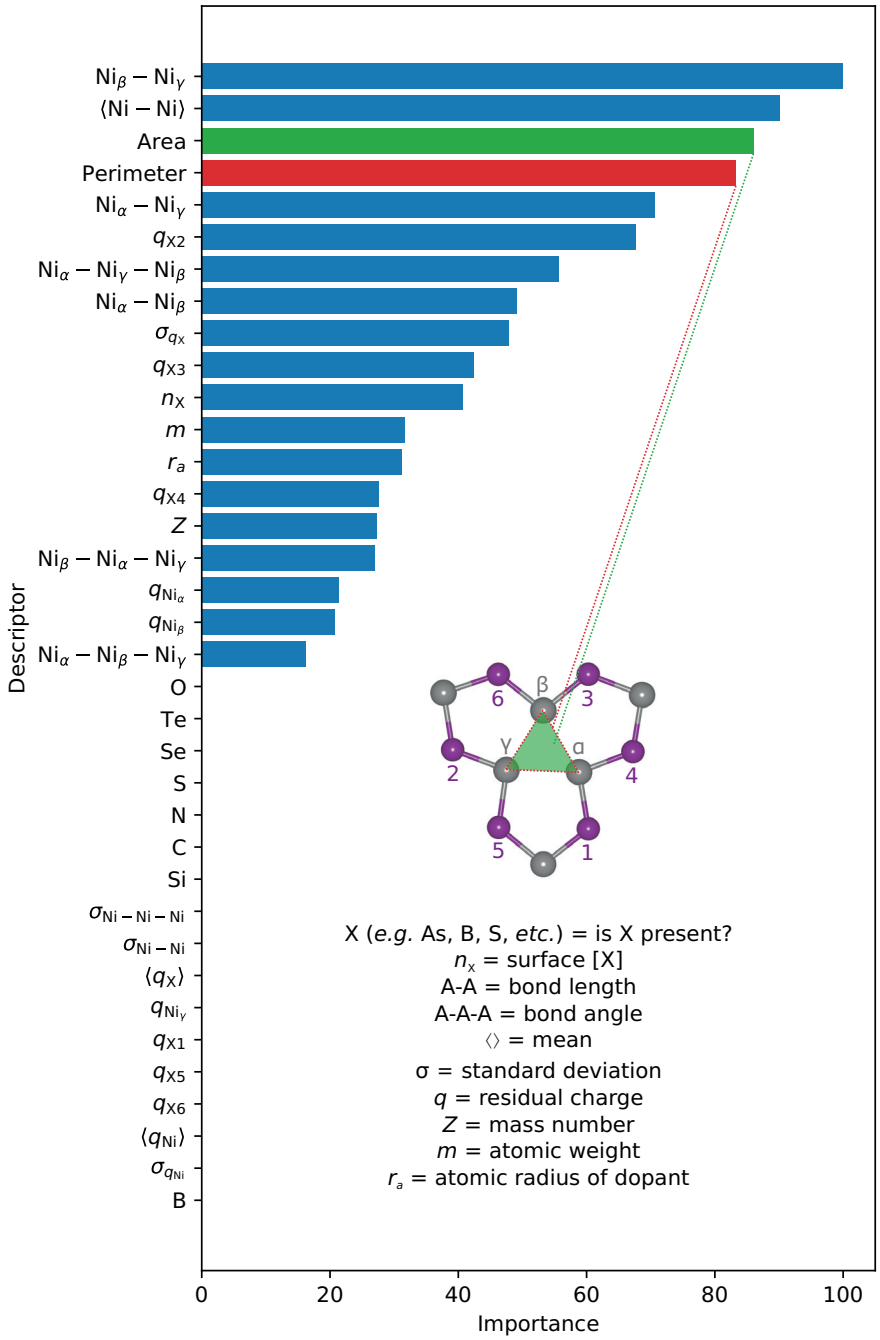

Figure C.2: Relative importance of descriptors calculated from RRF model. Inset shows definitions of descriptors. There are nine Ni atoms in the $(\sqrt{3} \times \sqrt{3})R30°$ $Ni_3P_2(0001)$ surface of Ni_2P. We only considered three Ni atoms (α, β, and γ) comprising one of the Ni_3-hollow sites for the descriptors.

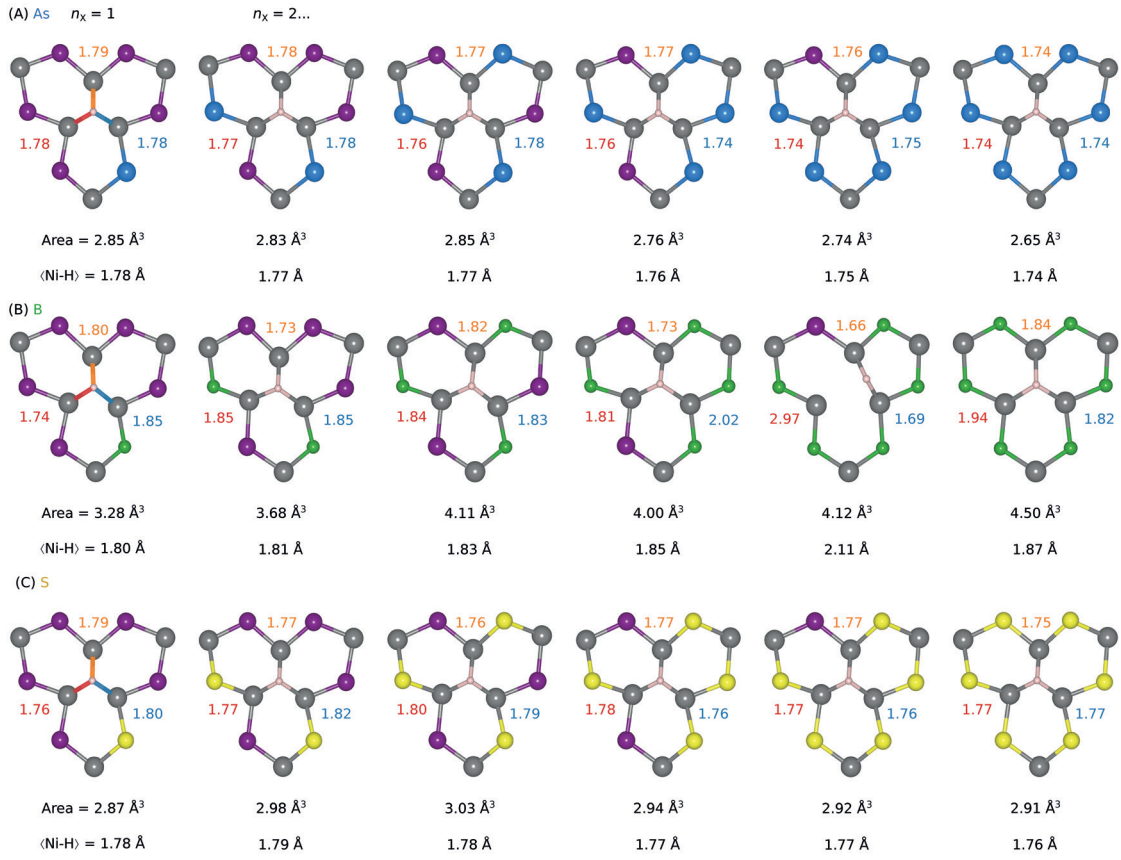

Figure C.3: Effect of Ni_3-hollow site expansion, induced by doping with nonmetals ($n_X = 1 - 6$), on the strength of Ni-H bonds. Only the Ni_3-hollow site and its local coordination environment are shown. As, B, and S atoms are blue, green, and yellow, respectively. The three unique Ni-H pairs and their bond lengths are colored blue, orange, and red. We used the area of the Ni_3-hollow site to measure its expansion and average Ni-H bond length ($\langle Ni - H \rangle$) to measure Ni-H bond strength. Although $\langle Ni - H \rangle$ appears to be a good measure of ΔG_H, we didn't include such descriptors involving explicitly H. (A) Doping with As induces minimal changes in the area of the Ni_3-hollow site and therefore its effect on Ni-H bond strength is very small. (B) Doping with B, in general, induces significant expansion of the Ni_3-hollow site thereby weakening the Ni-H bond. Additionally, at $n_B = 5$, one of the Ni-H bonds is severed. (C) Doping with S shows two regimes. For $n_S = 1 - 3$, doping induces expansion of the Ni_3-hollow, which consequently weakens Ni-H bond strength. For $n_S = 4 - 6$, however, doping induces compression of the Ni_3-hollow site causing the Ni-H bond to become stronger.

APPENDIX D : Supplemental: Surface crystal structure prediction using *ab initio* grand canonical Monte Carlo simulations

D.1. Additional Computational Details

D.1.1. DFT

We used a plane wave basis set with an energy cutoff of 50 Ry. We used Methfessel-Paxton first-order spreading (342) of 0.1 eV to improve self-consistent field convergence for metallic structures. k-point grids and energy cutoffs for bulk structures (Ag, Ag_2O, O_2) were selected such that the total energy was converged to within 3 meV/atom (see Table D.1). Calculated lattice constants are in good agreement with experimental values (see Table D.2). The number of layers in the slab model were selected such that the surface energy was converged to within 1 meV/Å^2.

Material	k-point grid	Energy cutoff (Ry)
Ag	$11 \times 11 \times 11$	50
Ag_2O	$4 \times 4 \times 4$	50
O_2	Γ	50

Table D.1: Converged k-point grids and energy cutoffs for bulk Ag, Ag_2O, and O_2

	Ag		Ag_2O	
	Calc.	Exp.	Calc.	Exp.
$a = b = c$	2.86	2.89	4.68	4.72

Table D.2: Lattice constants for bulk Ag and Ag_2O in Å.

D.1.2. Machine Learning

We removed highly correlated (Pearson correlation coefficient > 0.95) and near zero-variance features (variance < 0.05) from the data set. We randomly split the data set into a training and testing set containing 2/3 and 1/3 of the data, respectively. We then tuned the parameters of a random forest to maximize the R^2 of the testing set (see Table D.3).

Parameter	Definition	Value
`n_estimators`	# decision trees in random forest	40
`max_features`	# features to consider for each decision tree	11
`max_depth`	max depth of each decision tree	15

Table D.3: Optimal parameters of the random forest regressor implemented in scikit learn. (4)

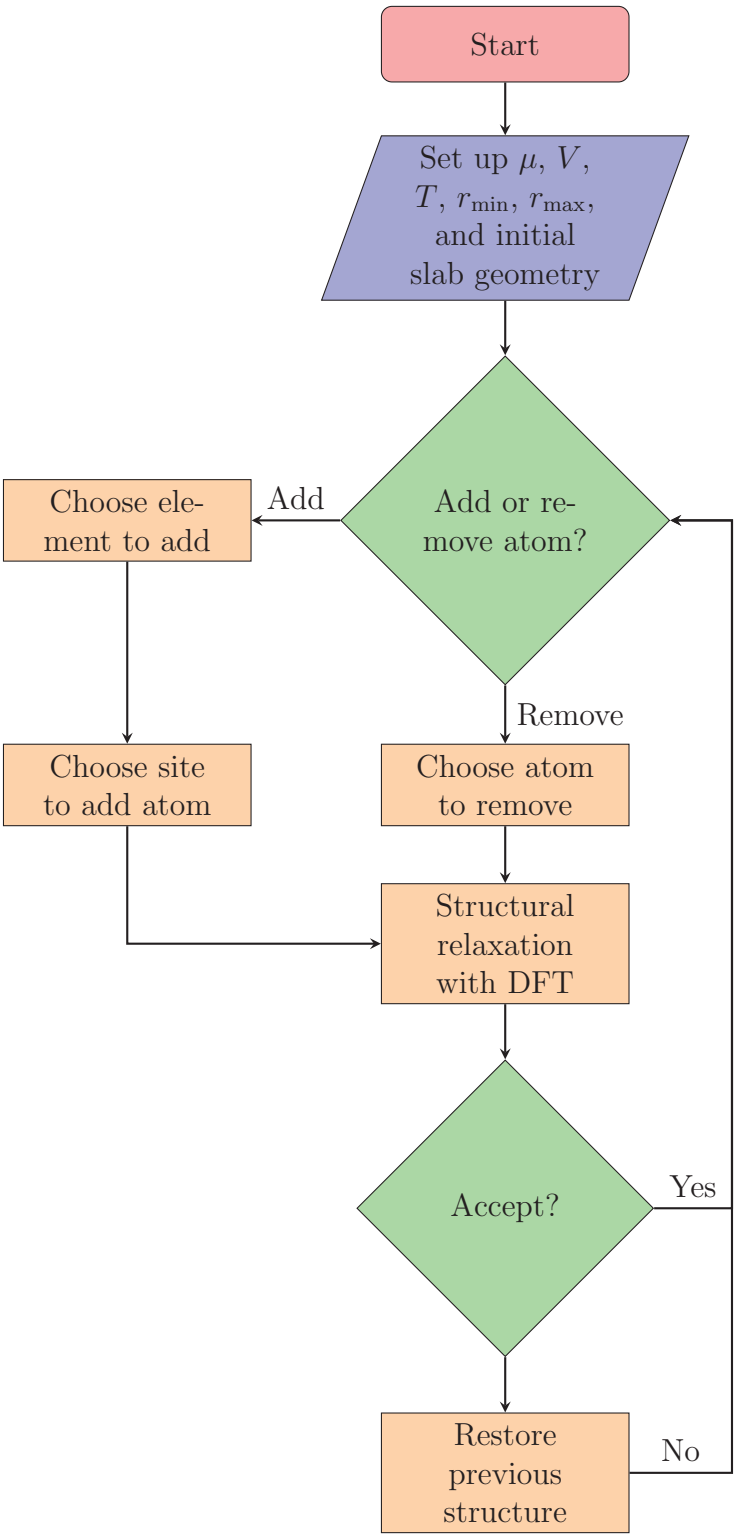

Figure D.1: Flowchart for *ab initio* GCMC

BIBLIOGRAPHY

[1] Wagman, D. D.; Evans, W. H.; Halow, I.; Parker, V. B.; Bailey, S. M.; Schumm, R. H. *Selected Values of Chemical Thermodynamic Properties*; National Bureau of Standards: Washington, D. C., 1965–1969.

[2] Blöchl, P. E. *Phys. Rev. B* **2000**, *62*, 6158.

[3] Martirez, J. M. P.; Carter, E. A. *ACS Nano* **2016**, *10*, 2940–2949.

[4] Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

[5] Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning*, 12th ed.; Springer Series in Statistics; Springer Science & Business Media, Inc.: 233 Spring St., New York, NY, 10013, USA, 2017.

[6] Deutschmann, O.; Knzinger, H.; Kochloefl, K.; Turek, T. *Heterogeneous Catalysis and Solid Catalysts, 3. Industrial Applications*; John Wiley & Sons, Ltd, 2011.

[7] Blaser, H.-U.; Pugin, B. *The Industrial Application of Heterogeneous Enantioselective Catalysts*; John Wiley & Sons, Ltd, 2008; Chapter 12, pp 413–437.

[8] Sheldon, R. A.; van Bekkum, H. *Fine Chemicals through Heterogeneous Catalysis*; John Wiley & Sons, Ltd, 2007.

[9] Romano, U.; Ricci, M. *Industrial Applications*; John Wiley & Sons, Ltd, 2013; Chapter 10, pp 451–506.

[10] Erisman, J. W.; Sutton, M. A.; Galloway, J.; Klimont, Z.; Winiwarter, W. *Nat. Geosci.* **2008**, *1*, 636.

[11] Honkala, K.; Hellman, A.; Remediakis, I. N.; Logadottir, A.; Carlsson, A.; Dahl, S.; Christensen, C. H.; Nørskov, J. K. *Science* **2005**, *307*, 555–8.

[12] Schlögl, R. *Angew. Chem. Int. Ed.* **2003**, *42*, 2004–2008.

[13] Xia, Q.-H.; Ge, H.-Q.; Ye, C.-P.; Liu, Z.-M.; Su, K.-X. *Chem. Rev.* **2005**, *105*, 1603–1662.

[14] Christopher, P.; Linic, S. *J. Am. Chem. Soc.* **2008**, *130*, 11264–11265.

[15] Oyama, S. T. In *Mechanisms in Homogeneous and Heterogeneous Epoxidation Catalysis*; Oyama, S. T., Ed.; Elsevier: Amsterdam, 2008; pp 3 – 99.

[16] Angelici, R. J. *Acc. Chem. Res.* **1988**, *21*, 387–394.

[17] Nørskov, J.; Clausen, B.; Topsøe, H. *Catal. Lett.* **1992**, *13*, 1–8.

[18] Cheng, F.; Chen, J.; Gou, X. *Adv. Mater.* **2006**, *18*, 2561–2564.

[19] Whipple, D. T.; Kenis, P. J. *J. Phys. Chem. Lett.* **2010**, *1*, 3451–3458.

[20] Kumar, B.; Brian, J. P.; Atla, V.; Kumari, S.; Bertram, K. A.; White, R. T.; Spurgeon, J. M. *Catal. Today* **2016**, *270*, 19–30.

[21] Zhang, L.; Zhao, Z.-J.; Gong, J. *Angewandte Chemie International Edition* **2017**, *56*, 11326–11353.

[22] Greeley, J.; Jaramillo, T. F.; Bonde, J.; Chorkendorff, I.; Norskov, J. K. *Computational high-throughput screening of electrocatalytic materials for hydrogen evolution*; World Scientific, 2011; pp 280–284.

[23] McKone, J. R.; Marinescu, S. C.; Brunschwig, B. S.; Winkler, J. R.; Gray, H. B. *Chem. Sci.* **2014**, *5*, 865–878.

[24] Zeng, M.; Li, Y. *J. Mater. Chem. A.* **2015**, *3*, 14942–14962.

[25] Jamesh, M. I. *J. Power Sources* **2016**, *333*, 213–236.

[26] Jiang, P.; Liu, Q.; Liang, Y.; Tian, J.; Asiri, A. M.; Sun, X. *Angewandte Chemie* **2014**, *126*, 13069–13073.

[27] Kibsgaard, J.; Tsai, C.; Chan, K.; Benck, J. D.; Nørskov, J. K.; Abild-Pedersen, F.; Jaramillo, T. F. *Energy Environ. Sci.* **2015**, *8*, 3022–3029.

[28] Xu, Y.; Wu, R.; Zhang, J.; Shi, Y.; Zhang, B. *ChemComm* **2013**, *49*, 6656–6658.

[29] Popczun, E. J.; Read, C. G.; Roske, C. W.; Lewis, N. S.; Schaak, R. E. *Angew. Chem. Int. Ed.* **2014**, *53*, 5427–5430.

[30] Liu, Q.; Tian, J.; Cui, W.; Jiang, P.; Cheng, N.; Asiri, A. M.; Sun, X. *Angew. Chem. Int. Ed.* **2014**, *53*, 6710–6714.

[31] Zhu, Y.-P.; Liu, Y.-P.; Ren, T.-Z.; Yuan, Z.-Y. *Adv. Funct. Mater.* **2015**, *25*, 7337–7347.

[32] Popczun, E. J.; McKone, J. R.; Read, C. G.; Biacchi, A. J.; Wiltrout, A. M.; Lewis, N. S.; Schaak, R. E. *J. Am. Chem. Soc.* **2013**, *135*, 9267–9270.

[33] Liu, P.; Rodriguez, J. A. *J. Am. Chem. Soc.* **2005**, *127*, 14871–14878.

[34] Laursen, A. B.; Patraju, K. R.; Whitaker, M. J.; Retuerto, M.; Sarkar, T.; Yao, N.; Ramanujachary, K. V.; Greenblatt, M.; Dismukes, G. C. *Energy Environ. Sci.* **2015**, *8*, 1027–1034.

[35] Stern, L.-A.; Feng, L.; Song, F.; Hu, X. *Energy Environ. Sci.* **2015**, *8*, 2347–2351.

[36] Hansen, M. H.; Stern, L.-A.; Feng, L.; Rossmeisl, J.; Hu, X. *Phys. Chem. Chem. Phys.* **2015**, *17*, 10823–10829.

[37] Moon, J.-S.; Jang, J.-H.; Kim, E.-G.; Chung, Y.-H.; Yoo, S. J.; Lee, Y.-K. *J. Catal.* **2015**, *326*, 92–99.

[38] Wexler, R. B.; Martirez, J. M. P.; Rappe, A. M. *ACS Catal.* **2017**, *7*, 7718–7725.

[39] Wexler, R. B.; Martirez, J. M. P.; Rappe, A. M. *J. Am. Chem. Soc.* **2018**, *140*, 4678–4683.

[40] Laursen, A. B.; Wexler, R. B.; Whitaker, M. J.; Izett, E. J.; Calvinho, K. U.; Hwang, S.; Rucker, R.; Wang, H.; Li, J.; Garfunkel, E.; et al. *ACS Catal.* **2018**, *8*, 4408–4419.

[41] Alerhand, O. L.; Vanderbilt, D.; Meade, R. D.; Joannopoulos, J. D. *Phys. Rev. Lett.* **1988**, *61*, 1973–6.

[42] Chadi, D. *Phys. Rev. Lett.* **1979**, *43*, 43.

[43] Ramstad, A.; Brocks, G.; Kelly, P. *Phys. Rev. B.* **1995**, *51*, 14504.

[44] Giessibl, F. J. *Science* **1995**, *267*, 68–71.

[45] Binnig, G.; Rohrer, H.; Gerber, C.; Weibel, E. *Phys. Rev. Lett.* **1983**, *50*, 120.

[46] Takayanagi, K.; Tanishiro, Y.; Takahashi, S.; Takahashi, M. *Surf. Sci.* **1985**, *164*, 367–392.

[47] Tersoff, J.; Hamann, D. *Phys. Rev. Lett.* **1983**, *50*, 1998.

[48] Binnig, G.; Rohrer, H.; Gerber, C.; Weibel, E. *Phys. Rev. Lett.* **1982**, *49*, 57.

[49] Kolpak, A. M.; Li, D.; Shao, R.; Rappe, A. M.; Bonnell, D. A. *Phys. Rev. Lett.* **2008**, *101*, 036102 (1–4).

[50] Saidi, W. A.; Martirez, J. M. P.; Rappe, A. M. *Nano Lett.* **2014**, *14*, 6711–6717.

[51] Martirez, J. M. P.; Kim, S.; Morales, E. H.; Diroll, B. T.; Cargnello, M.; Gordon, T. R.; Murray, C. B.; Bonnell, D. A.; Rappe, A. M. *J. Am. Chem. Soc.* **2015**, *173*, 2939–2947.

[52] Stampfl, C.; Ganduglia-Pirovano, M. V.; Reuter, K.; Scheffler, M. *Surf. Sci.* **2002**, *500*, 368–394.

[53] Reuter, K.; Stampf, C.; Scheffler, M. *Ab Initio Atomistic Thermodynamics and Statistical Mechanics of Surface Properties and Functions*; Springer, 2008; pp 149–194.

[54] Rong, X.; Kolpak, A. M. *J. Phys. Chem. Lett.* **2015**, *6*, 1785–1789.

[55] Carlisle, C. I.; andM. L. Bocquet, D. A. K.; Cerdá, J.; Sautet, P. *Phys. Rev. Lett.* **2000**, *84*, 3899–902.

[56] Schnadt, J.; Michaelides, A.; Knudsen, J.; Vang, R. T.; Reuter, K.; Lægsgaard, E.; Scheffler, M.; Besenbacher, F. *Phys. Rev. Lett.* **2006**, *96*, 146101.

[57] Reuter, K.; Scheffler, M. *Phys. Rev. B* **2001**, *65*, 035406–1–11.

[58] Levchenko, S. V.; Rappe, A. M. *Phys. Rev. Lett.* **2008**, *100*, 256101 (1–4).

[59] Oganov, A. R.; Glass, C. W. *J. Chem. Phys.* **2006**, *124*, 244704.

[60] Lyakhov, A. O.; Oganov, A. R.; Stokes, H. T.; Zhu, Q. *Comput. Phys. Commun.* **2013**, *184*, 1172–1182.

[61] Oganov, A. R.; Lyakhov, A. O.; Valle, M. *Acc. Chem. Res.* **2011**, *44*, 227–237.

[62] Zhu, Q.; Li, L.; Oganov, A. R.; Allen, P. B. *Phys. Rev. B.* **2013**, *87*, 195317.

[63] Wang, Y.; Lv, J.; Zhu, L.; Ma, Y. *Phys. Rev. B.* **2010**, *82*, 094116.

[64] Wang, Y.; Lv, J.; Zhu, L.; Ma, Y. *Comput. Phys. Commun.* **2012**, *183*, 2063–2070.

[65] Lu, S.; Wang, Y.; Liu, H.; Miao, M.-s.; Ma, Y. *Nat. Commun.* **2014**, *5*, 3666.

[66] Ulissi, Z. W.; Singh, A. R.; Tsai, C.; Nørskov, J. K. *J. Chem. Phys. Lett.* **2016**, *7*, 3931–3935.

[67] Wexler, R. B.; Martirez, J. M. P.; Rappe, A. M. *Chem. Mater.* **2016**, *28*, 5365–5372.

[68] Wexler, R. B.; Qiu, T.; Rappe, A. M. *J. Phys. Chem. C.* **2019**, *123*, 2321–2328.

[69] Born, M.; Oppenheimer, J. R. *Ann. Phys.* **1927**, *84*, 457–484.

[70] Schrödinger, E. *Phys. Rev.* **1926**, *28*, 1049–1070.

[71] Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864.

[72] Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133–A1138.

[73] Schwartz, S. D. *Theoretical Methods in Condensed Phase Chemistry*; Springer Science & Business Media, 2002; Vol. 5.

[74] Wesolowski, T. A.; Wang, Y. A. *Recent Progress in Orbital-Free Density Functional Theory*; World Scientific, 2013; Vol. 6.

[75] Bach, V.; Delle Site, L. *Many-Electron Approaches in Physics, Chemistry and Mathematics: A Multidisciplinary View*; Springer, 2014.

[76] Perdew, J. P.; Zunger, A. *Phys. Rev. B* **1981**, *23*, 5048–79.

[77] Ceperley, D. M.; Alder, B. J. *Phys. Rev. Lett.* **1980**, *45*, 566–569.

[78] Perdew, J. P.; Wang, Y. *Phys. Rev. B* **1992**, *45*, 13244–9.

[79] Kohn, W.; Becke, A. D.; Parr, R. G. *J. Phys. Chem.* **1996**, *100*, 12974–12980.

[80] Perdew, J. P.; Chevary, J. A.; Vosko, S. H.; Jackson, K. A.; Pederson, M. R.; Singh, D. J.; Fiolhais, C. *Phys. Rev. B* **1992**, *46*, 6671–6687.

[81] Becke, A. D. *J. Chem. Phys.* **1992**, *96*, 2155–2160.

[82] Proynov, E.; Ruiz, E.; Vela, A.; Salahub, D. *Int. J. Quantum Chem.* **1995**, *56*, 61–78.

[83] Hammer, B.; Jacobsen, K. W.; Nørskov, J. K. *Phys. Rev. Lett.* **1993**, *70*, 3971.

[84] Hammer, B.; Scheffler, M. *Phys. Rev. Lett.* **1995**, *74*, 3487.

[85] Hamann, D. R. *Phys. Rev. Lett.* **1996**, *76*, 660–3.

[86] Philipsen, P.; te Velde, G.; Baerends, E. *Chem. Phys. Lett.* **1994**, *226*, 583–588.

[87] Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865 (1–8).

[88] Wang, Y.; Perdew, J. P. *Phys. Rev. B.* **1991**, *43*, 8911.

[89] Levy, M. *Int. J. Quantum Chem.* **1989**, *36*, 617–619.

[90] Oliver, G.; Perdew, J. *Phys. Rev. A.* **1979**, *20*, 397.

[91] Bowen, C.; Sugiyama, G.; Alder, B. *Phys. Rev. B.* **1994**, *50*, 14838.

[92] Moroni, S.; Ceperley, D. M.; Senatore, G. *Phys. Rev. Lett.* **1995**, *75*, 689.

[93] Ortiz, G. *Phys. Rev. B.* **1992**, *45*, 11328.

[94] Becke, A. D. *J. Chem. Phys.* **1986**, *84*, 4524–4529.

[95] Bloch, F. *Z. Physik* **1928**, *52*, 555.

[96] Chadi, D. J.; Cohen, M. L. *Phys. Rev. B* **1973**, *8*, 5747–5753.

[97] Monkhorst, H. J.; Pack, J. D. *Phys. Rev. B* **1976**, *13*, 5188–5192.

[98] Heitler, W.; London, F. *Z. Phys.* **1927**, *44*, 455–472.

[99] Pauling, L. *The Nature of the Chemical Bond*, 3rd ed.; Cornell University Press: Ithaca, NY, 1960.

[100] Slater, J. C. *Phys. Rev.* **1930**, *35*, 509.

[101] Wheland, G. W. *Resonance in Organic Chemistry*; John Wiley & Sons, 1955.

[102] Eyring, H.; Polanyi, M. *Z. Phys. Chem.* **1931**, *B12*, 279–311.

[103] Hamann, D. R.; Schlüter, M.; Chiang, C. *Phys. Rev. Lett.* **1979**, *43*, 1494–7.

[104] Kleinman, L.; Bylander, D. M. *Phys. Rev. Lett.* **1982**, *48*, 1425–8.

[105] Ramer, N. J.; Rappe, A. M. *Phys. Rev. B* **1999**, *59*, 12471–12478.

[106] Rappe, A. M.; Rabe, K. M.; Kaxiras, E.; Joannopoulos, J. D. *Phys. Rev. B Rapid Comm.* **1990**, *41*, 1227–1230.

[107] Allen, M. J.; Tozer, D. J. *J. Chem. Phys.* **2002**, *117*, 11113–11120.

[108] Kristyán, S.; Pulay, P. *Chem. Phys. Lett.* **1994**, *229*, 175–180.

[109] Slavıček, P.; Kalus, R.; Paška, P.; Odvárková, I.; Hobza, P.; Malijevskỳ, A. *J. Chem. Phys.* **2003**, *119*, 2102–2119.

[110] Sherrill, C. D.; Takatani, T.; Hohenstein, E. G. *J. Phys. Chem. A.* **2009**, *113*, 10146–10159.

[111] Goerigk, L.; Grimme, S. *J. Chem. Theory Comput.* **2009**, *6*, 107–126.

[112] Goerigk, L.; Grimme, S. *J. Chem. Theory Comput.* **2011**, *7*, 291–309.

[113] Huenerbein, R.; Schirmer, B.; Moellmann, J.; Grimme, S. *Phys. Chem. Chem. Phys.* **2010**, *12*, 6940–6948.

[114] Tao, J.; Perdew, J. P.; Staroverov, V. N.; Scuseria, G. E. *Phys. Rev. Lett.* **2003**, *91*, 146401.

[115] Zhao, Y.; Truhlar, D. G. *J. Phys. Chem. A.* **2005**, *109*, 5656–5667.

[116] Grimme, S. *J. Chem. Phys.* **2006**, *124*, 034108.

[117] Karton, A.; Tarnopolsky, A.; Lamére, J.-F.; Schatz, G. C.; Martin, J. M. *J. Phys. Chem. A.* **2008**, *112*, 12868–12886.

[118] Schwabe, T.; Grimme, S.; Djukic, J.-P. *J. Am. Chem. Soc.* **2009**, *131*, 14156–14157.

[119] Grimme, S.; Djukic, J.-P. *Inorg. Chem.* **2010**, *49*, 2911–2919.

[120] Dion, M.; Rydberg, H.; Schröder, E.; Langreth, D. C.; Lundqvist, B. I. *Phys. Rev. Lett.* **2004**, *92*, 246401–1–4.

[121] Lee, K.; Murray, E. D.; Kong, L.; Lundqvist, B. I.; Langreth, D. C. *Phys. Rev. B* **2010**, *82*, 081101.

[122] Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157–167.

[123] Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104.

[124] Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463–1473.

[125] Jurečka, P.; Černỳ, J.; Hobza, P.; Salahub, D. R. *J. Comput. Chem.* **2007**, *28*, 555–569.

[126] von Lilienfeld, O. A.; Tavernelli, I.; Rothlisberger, U.; Sebastiani, D. *Phys. Rev. Lett.* **2004**, *93*, 153004.

[127] Sun, Y.; Kim, Y.-H.; Lee, K.; Zhang, S. *J. Chem. Phys.* **2008**, *129*, 154102.

[128] Johnson, E. R.; Mackie, I. D.; DiLabio, G. A. *J. Phys. Org. Chem.* **2009**, *22*, 1127–1135.

[129] Grimme, S. *J. Comput. Chem.* **2006**, *27*, 1787–1799.

[130] Eisenschitz, R.; London, F. *Z. Phys.* **1930**, *60*, 491–527.

[131] London, F. *Z. Phys.* **1930**, *63*, 245–279.

[132] London, F. *Trans. Faraday Soc.* **1937**, *33*, 8–26.

[133] Chase Jr, M. W. *J. Phys. Chem. Ref. Data, Monograph* **1998**, *9*, 1–1951.

[134] Barin, I. *Thermochemical Data of Pure Substances*; Federal Republic of Germany, 1989.

[135] Huber, K.-P. *Molecular Spectra and Molecular Structure: IV. Constants of Diatomic Molecules*; Springer Science & Business Media, 2013.

[136] Baro, A.; Ibach, H.; Bruchmann, H. *Surf. Sci.* **1979**, *88*, 384–398.

[137] Baroni, S.; Giannozzi, P.; Testa, A. *Phys. Rev. Lett.* **1987**, *58*, 1861–64.

[138] PAULING, L. *J. Am. Chem. Soc.* **1929**, *51*, 1010–26.

[139] Van Vleck, J. *Phys. Rev.* **1932**, *41*, 208.

[140] Griffith, J.; Orgel, L. *Q. Rev. Chem. Soc.* **1957**, *11*, 381–393.

[141] Tasker, P. W. *Journal of Physics C: Solid State Physics* **1979**, *12*, 4977–4984.

[142] Pourbaix, M. *Atlas of Electrochemical Equilibria in Aqueous Solution*, 2nd ed.; National Association of Corrosion, 1974.

[143] Irikura, K. K. *J. Phys. Chem. Ref. Data* **2007**, *36*, 389–397.

[144] Shimanouchi, T. *Tables of Molecular Vibrational Frequencies*; National Bureau of Standards: Washington, DC, 1980; Vol. 1.

[145] Metropolis, N.; Rosenbluth, A.; Rosenbluth, M.; Teller, A.; Teller, M. *Journal of Chemical Physics* **1953**, *21*, 1087–1092.

[146] Henkelman, G.; Arnaldsson, A.; Jónsson, H. *Comput. Mater. Sci.* **2006**, *36*, 354–360.

[147] Sanville, E.; Kenny, S. D.; Smith, R.; Henkelman, G. *J. Comp. Chem.* **2007**, *28*, 899–908.

[148] Tang, W.; Sanville, E.; Henkelman, G. *J. Phys. Condens. Matter.* **2009**, *21*, 084204.

[149] Yu, M.; Trinkle, D. R. *J. Chem. Phys.* **2011**, *134*, 064111.

[150] Hammer, B.; Norskov, J. *Nature* **1995**, *376*, 238–240.

[151] Hammer, B.; Nørskov, J. K. In *Advances in Catalysis*; Elsevier, 2000; Vol. 45, pp 71–129.

[152] Mulliken, R. S. *J. Chem. Phys.* **1955**, *23*, 1833.

[153] Kuhn, M.; et al. *J. Stat. Softw.* **2008**, *28*, 1–26.

[154] Pedregosa, F.; et al. *J. Mach. Learn. Res.* **2011**, *12*, 2825–2830.

[155] Breiman, L.; Friedman, J.; Stone, C.; Olshen, R. *Classification and Regression Trees*; Taylor & Francis, 1984.

[156] Ripley, B. *Pattern Recognition and Neural Networks*; Cambridge University Press, 1996.

[157] Deng, H.; Runger, G. In *The 2012 International Joint Conference on Neural Networks (IJCNN)*; IEEE; pp 1–8.

[158] Deng, H. *CoRR* **2013**, *abs/1306.0237*.

[159] Deng, H.; Runger, G. *Pattern Recognit.* **2013**, *46*, 3483–3489.

[160] Hastie, T.; Tibshirani, R.; Friedman, J. *The Elements of Statistical Learning: Data Mining, Inference, and Prediction*; Springer Series in Statistics; Springer: New York, 2013.

[161] Haga, K.; Shiratori, Y.; Nojiri, Y.; Ito, K.; Sasaki, K. *J. Electrochem. Soc.* **2010**, *157*, B1693–B1700.

[162] Colletti, L. P.; Teklay, D.; Stickney, J. L. *J. Electroanal. Chem.* **1994**, *369*, 145–152.

[163] Meissner, D.; Benndorf, C.; Memming, R. *Appl. Surf. Sci.* **1987**, *27*, 423–436.

[164] Oyama, S. T. *J. Catal.* **2003**, *216*, 343–352.

[165] Liu, P.; Rodriguez, J. A.; Takahashi, Y.; Nakamura, K. *J. Catal.* **2009**, *262*, 294–303.

[166] Liu, X.; Chen, J.; Zhang, J. *Ind. Eng. Chem. Res.* **2008**, *47*, 5362–5368.

[167] Aronsson, B.; Lundstrøm, T.; Rundqvist, S. *Borides, Silicides, and Phosphides: A Critical Review of Their Preparation, Properties, and Crystal Chemistry*; Methuen: London, England, 1965.

[168] Corbridge, D. E. C. *Studies in Inorganic Chemistry*, 4th ed.; Elsevier: Amsterdam, Netherlands, 1990; Vol. 10.

[169] Edamoto, K.; Nakadai, Y.; Inomata, H.; Ozawa, K.; Otani, S. *Solid State Commun.* **2008**, *148*, 135–138.

[170] Falch, S.; Lamparter, P.; Steeb, S. *Z. Naturforsch.* **1984**, *39a*, 1175–1183.

[171] Kanama, D.; Oyama, S. T.; Otani, S.; Cox, D. F. *Surf. Sci.* **2004**, *552*, 8–16.

[172] Li, Q.; Hu, X. *Phys. Rev. B.* **2006**, *74*, 035414 (1–5).

[173] Moula, M. G.; Suzuki, S.; Chun, W.-J.; Omani, S.; Oyama, S. T.; Sakura, K. *Surf. Interface Anal.* **2006**, *38*, 1611–1614.

[174] Suzuki, S.; Moula, G. M.; Miyamoto, T.; Nakagawa, Y.; Kinoshita, K.; Asakura, K.; Oyama, S. T.; Omani, S. *J. Nanosci. Nanotechnol.* **2009**, *9*, 195–201.

[175] Hernandez, A. B.; Ariga, H.; Takakusagi, S.; Kinoshita, K.; Suzuki, S.; Otani, S.; Oyama, S. T.; Asakura, K. *Chem. Phys. Lett.* **2011**, *513*, 48–52.

[176] Kinoshita, K.; Simon, G. H.; König, T.; Heyde, M.; Freund, H.-J.; Nakagawa, Y.; Suzuki, S.; Chun, W.-J.; Oyama, S. T.; Omani, S.; Asakura, K. *Jpn. J. Appl. Phys.* **2008**, *47*, 6088.

[177] Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864–B871.

[178] Giannozzi, P.; et al. *J. Phys. Condens. Matter.* **2009**, *21*, 395502.

[179] http://opium.sourceforge.net.

[180] Bengtsson, L. *Phys. Rev. B* **1999**, *59*, 12301–12304.

[181] Larsson, E. *Arkiv Kemi* **1965**, *23*, 335–365.

[182] Guo, D.; Nakagawa, Y.; Ariga, H.; Suzuki, S.; Kinoshita, K.; Miyamoto, T.; Takakusagi, S.; Asakura, K.; Omani, S.; Oyama, S. T. *Surf. Sci.* **2010**, *604*, 1347–1352.

[183] Ariga, H.; Kawashima, M.; Takakusagi, S.; Asakura, K. *Chem. Lett.* **2013**, *42*, 1481–1483.

[184] Liu, P.; Rodriguez, J. A.; Asakura, T.; Gomes, J.; Nakamura, K. *J. Phys. Chem. B.* **2005**, *109*, 4575–4583.

[185] Kanama, D.; Oyama, S. T.; Otani, S.; Cox, D. F. *Surf. Sci. Spectra* **2001**, *8*, 220–224.

[186] Moon, J. S.; Kim, E. G.; Lee, Y. K. *J. Catal.* **2014**, *311*, 144–152.

[187] Maxwell, L. R.; Hendricks, S. B.; Mosley, V. M. *J. Chem. Phys.* **1935**, *3*, 699–709.

[188] Hellwege, K. H.; Hellwege, A. M. *Atomic and Molecular Physics*; Landolt-Bornstein, 1976.

[189] Kim, S.; Sinai, O.; Lee, C.-W.; Rappe, A. M. *Phys. Rev. B* **2015**, *92*, 235431 (1–8).

[190] Morales, E. H.; Martirez, J. M. P.; Saidi, W. A.; Rappe, A. M.; Bonnell, D. A. *ACS Nano* **2014**, *8*, 4465–4473.

[191] Martirez, J. M. P.; Morales, E. H.; Al-Saidi, W. A.; Bonnell, D. A.; Rappe, A. M. *Phys. Rev. Lett.* **2012**, *109*, 256802 (1–5).

[192] Marković, N.; Ross, P. N. *Surf. Sci. Rep.* **2002**, *45*, 117–229.

[193] Hinnemann, B.; Moses, P. G.; Bonde, J.; Jørgensen, K. P.; Nielsen, J. H.; Horch, S.; Chorkendorff, I.; Nørskov, J. K. *J. Am. Chem. Soc.* **2005**, *127*, 5308–5309.

[194] Greeley, J.; Jaramillo, T. F.; Bonde, J.; Chorkendorff, I.; Nørskov, J. K. *Nat. Mater.* **2006**, *5*, 909–913.

[195] de Chialvo, M. G.; Chialvo, A. *J. Electrochem. Soc.* **2000**, *147*, 1619–1622.

[196] de Chialvo, M. R. G.; Chialvo, A. C. *Phys. Chem. Chem. Phys.* **2004**, *6*, 4009–4017.

[197] Quaino, P. M.; de Chialvo, M. R. G.; Chialvo, A. C. *Phys. Chem. Chem. Phys.* **2004**, *6*, 4450–4455.

[198] Conway, B.; Bai, L. *J. Electroanal. Chem. Interfacial Electrochem.* **1986**, *198*, 149–175.

[199] Conway, B.; Bai, L. *Int. J. Hydrogen Energy* **1986**, *11*, 533–540.

[200] Zheng, Y.; Jiao, Y.; Jaroniec, M.; Qiao, S. Z. *Angew. Chem. Int. Ed.* **2015**, *54*, 52–65.

[201] Conway, B.; Tilak, B. *Adv. Catal.* **1992**, *38*, 1–147.

[202] Jaramillo, T. F.; Jørgensen, K. P.; Bonde, J.; Nielsen, J. H.; Horch, S.; Chorkendorff, I. *Science* **2007**, *317*, 100–102.

[203] Jaramillo, T. F.; Bonde, J.; Zhang, J.; Ooi, B.-L.; Andersson, K.; Ulstrup, J.; Chorkendorff, I. *J. Phys. Chem. C* **2008**, *112*, 17492–17498.

[204] Hou, Y.; Abrams, B. L.; Vesborg, P. C. K.; Björketun, M. E.; Herbst, K.; Bech, L.; Setti, A. M.; Damsgaard, C. D.; Pedersen, T.; Hansen, O.; Rossmeisl, J.; Dahl, S.; Nørskov, J. K.; Chorkendorff, I. *Nat. Mater.* **2011**, *10*, 434–438.

[205] Seger, B.; Herbst, K.; Pedersen, T.; Abrams, B.; Vesborg, P. C. K.; Hansen, O.; Chorkendorff, I. *J. Electrochem. Soc.* **2014**, *161*, H722–H724.

[206] Kibsgaard, J.; Jaramillo, T. F.; Besenbacher, F. *Nat. Chem.* **2014**, *6*, 248–253.

[207] Merki, D.; Fierro, S.; Vrubel, H.; Hu, X. *Chem. Sci.* **2011**, *2*, 1262–1267.

[208] Kibsgaard, J.; Chen, Z.; Reinecke, B. N.; Jaramillo, T. F. *Nat. Mater.* **2012**, *11*, 963–969.

[209] Li, Y.; Wang, H.; Xie, L.; Liang, Y.; Hong, G.; Dai, H. *J. Am. Chem. Soc.* **2011**, *133*, 7296–7299.

[210] Li, D. J.; Maiti, U. N.; Lim, J.; Choi, D. S.; Lee, W. J.; Oh, Y.; Lee, G. Y.; Kim, S. O. *Nano Lett.* **2014**, *14*, 1228–1233.

[211] Bonde, J.; Moses, P. G.; Jaramillo, T. F.; Nørskov, J. K.; Chorkendorff, I. *Faraday Discuss.* **2009**, *140*, 219–231.

[212] Merki, D.; Vrubel, H.; Rovelli, L.; Fierro, S.; Hu, X. *Chem. Sci.* **2012**, *3*, 2515–2525.

[213] Lukowski, M. A.; Daniel, A. S.; Meng, F.; Forticaux, A.; Li, L.; Jin, S. *J. Am. Chem. Soc.* **2013**, *135*, 10274–10277.

[214] Voiry, D.; Yamaguchi, H.; Li, J.; Silva, R.; Alves, D. C.; Fujita, T.; Chen, M.; Asefa, T.; Shenoy, V. B.; Eda, G.; Chhowalla, M. *Nat. Mater.* **2013**, *12*, 850–855.

[215] Wang, H.; Lu, Z.; Kong, D.; Sun, J.; Hymel, T. M.; Cui, Y. *ACS Nano* **2014**, *8*, 4940–4947.

[216] Vrubel, H.; Hu, X. *Angew. Chem.* **2012**, *124*, 12875–12878.

[217] Chen, W.-F.; Wang, C.-H.; Sasaki, K.; Marinkovic, N.; Xu, W.; Muckerman, J.; Zhu, Y.; Adzic, R. *Energy Environ Sci.* **2013**, *6*, 943–951.

[218] Esposito, D. V.; Hunt, S. T.; Stottlemyer, A. L.; Dobson, K. D.; McCandless, B. E.; Birkmire, R. W.; Chen, J. G. *Angew. Chem. Int. Ed.* **2010**, *49*, 9859–9862.

[219] Wu, H. B.; Xia, B. Y.; Yu, L.; Yu, X.-Y.; Lou, X. W. D. *Nat. Commun.* **2015**, *6*, 1–8.

[220] Esposito, D. V.; Hunt, S. T.; Kimmel, Y. C.; Chen, J. G. *J. Am. Chem. Soc* **2012**, *134*, 3025–3033.

[221] Youn, D. H.; Han, S.; Kim, J. Y.; Kim, J. Y.; Park, H.; Choi, S. H.; Lee, J. S. *ACS Nano* **2014**, *8*, 5164–5173.

[222] Wan, C.; Regmi, Y. N.; Leonard, B. M. *Angew. Chem.* **2014**, *126*, 6525–6528.

[223] Esposito, D. V.; Chen, J. G. *Energy Environ Sci.* **2011**, *4*, 3900–3912.

[224] Harnisch, F.; Sievers, G.; Schröder, U. *Appl. Catal., B* **2009**, *89*, 455–458.

[225] Yan, Y.; Xia, B.; Qi, X.; Wang, H.; Xu, R.; Wang, J.-Y.; Zhang, H.; Wang, X. *Chem. Commun.* **2013**, *49*, 4884–4886.

[226] Liu, Y.; Yu, G.; Li, G.-D.; Sun, Y.; Asefa, T.; Chen, W.; Zou, X. *Angew. Chem. Int. Ed.* **2015**, *54*, 10752–10757.

[227] Chen, W.-F.; Sasaki, K.; Ma, C.; Frenkel, A. I.; Marinkovic, N.; Muckerman, J. T.; Zhu, Y.; Adzic, R. R. *Angew. Chem. Int. Ed.* **2012**, *51*, 6131–6135.

[228] Zheng, P.; Zhao, J.; Zheng, J.; Ma, G.; Zhu, Z. *J. Mater. Chem.* **2012**, *22*, 12116–12120.

[229] Su, J.; Lu, N.; Zhao, J.; Yu, H.; Huang, H.; Dong, X.; Quan, X. *J. Hazard. Mater.* **2012**, *231*, 105–113.

[230] Ishikawa, A.; Takata, T.; Kondo, J. N.; Hara, M.; Domen, K. *J. Phys. Chem. B* **2004**, *108*, 11049–11053.

[231] Wirth, S.; Harnisch, F.; Weinmann, M.; Schröder, U. *Appl. Catal., B* **2012**, *126*, 225–230.

[232] Choi, D.; Blomgren, G. E.; Kumta, P. N. *Adv. Mater.* **2006**, *18*, 1178–1182.

[233] Choi, D.; Kumta, P. N. *Electrochem. Solid State Lett.* **2005**, *8*, A418–A422.

[234] Pan, Y.; Liu, Y.; Zhao, J.; Yang, K.; Liang, J.; Liu, D.; Hu, W.; Liu, D.; Liu, Y.; Liu, C. *J. Mater. Chem. A.* **2015**, *3*, 1656–1665.

[235] Koocher, N. Z.; Martirez, J. M. P.; Rappe, A. M. *J. Phys. Chem. Lett.* **2014**, *5*, 3408–3414.

[236] Kim, S.; Schoenberg, M. R.; Rappe, A. M. *Phys. Rev. Lett.* **2011**, *107*, 076102 (1–5).

[237] Kolpak, A. M.; Grinberg, I.; Rappe, A. M. *Phys. Rev. Lett.* **2007**, *98*, 166101 (1–4).

[238] Kucernak, A. R.; Sundaram, V. N. N. *J. Mater. Chem. A* **2014**, *2*, 17435–17445.

[239] Jin, L.; Xia, H.; Huang, Z.; Lv, C.; Wang, J.; Humphrey, M. G.; Zhang, C. *J. Mater. Chem. A* **2016**, *4*, 10925–10932.

[240] Giannozzi, P.; et al. *J. Phys.: Condens. Matter* **2009**, *21*, 395502 (1–19).

[241] Ramalho, J. P. P.; Gomes, J. R.; Illas, F. *RSC Advances* **2013**, *3*, 13085–13100.

[242] Eder, F.; Lercher, J. A. *Zeolites* **1997**, *18*, 75–81.

[243] Chizmeshya, A.; Ritter, C.; Tolle, J.; Cook, C.; Menendez, J.; Kouvetakis, J. *Chem. Mater.* **2006**, *18*, 6266–6277.

[244] Cottrell, T. L. *The Strengths of Chemical Bonds, 2nd ed.*; Academic Press: London, 1958.

[245] Ledendecker, M.; Mondschein, J.; Kasian, O.; Geiger, S.; Göhl, D.; Schalenbach, M.; Zeradjanin, A.; Cherevko, S.; Schaak, R. E.; Mayrhofer, K. *Angew. Chem.* **2017**, *129*, 9899–9903.

[246] Kibsgaard, J.; Jaramillo, T. F. *Angew. Chem. Int. Ed.* **2014**, *53*, 14433–14437.

[247] Cabán-Acevedo, M.; Stone, M. L.; Schmidt, J.; Thomas, J. G.; Ding, Q.; Chang, H.-C.; Tsai, M.-L.; He, J.-H.; Jin, S. *Nat. Mater.* **2015**, *14*, 1245–1251.

[248] Luo, J.; Wang, H.; Su, G.; Tang, Y.; Liu, H.; Tian, F.; Li, D. *J. Mater. Chem. A* **2017**, *5*, 14865–14872.

[249] Feng, L.; Vrubel, H.; Bensimon, M.; Hu, X. *Phys. Chem. Chem. Phys.* **2014**, *16*, 5917–5921.

[250] Pu, Z.; Liu, Q.; Tang, C.; Asiri, A. M.; Sun, X. *Nanoscale* **2014**, *6*, 11031–11034.

[251] Pan, Y.; Hu, W.; Liu, D.; Liu, Y.; Liu, C. *J. Mater. Chem. A* **2015**, *3*, 13087–13094.

[252] Liang, H.; Gandi, A. N.; Anjum, D. H.; Wang, X.; Schwingenschlögl, U.; Alshareef, H. N. *Nano Lett* **2016**, *16*, 7718–7725.

[253] Feng, Y.; Yu, X.-Y.; Paik, U. *ChemComm* **2016**, *52*, 1633–1636.

[254] Li, J.; Yan, M.; Zhou, X.; Huang, Z.-Q.; Xia, Z.; Chang, C.-R.; Ma, Y.; Qu, Y. *Adv. Funct. Mater.* **2016**, *26*, 6785–6796.

[255] Wang, P.; Pu, Z.; Li, Y.; Wu, L.; Tu, Z.; Jiang, M.; Kou, Z.; Amiinu, I. S.; Mu, S. *ACS Appl. Mater. Interfaces* **2017**, *9*, 26001–26007.

[256] Zhang, Y.; Liu, Y.; Ma, M.; Ren, X.; Liu, Z.; Du, G.; Asiri, A. M.; Sun, X. *ChemComm* **2017**, *53*, 11048–11051.

[257] Sun, Y.; Hang, L.; Shen, Q.; Zhang, T.; Li, H.; Zhang, X.; Lyu, X.; Li, Y. *Nanoscale* **2017**, *9*, 16674–16679.

[258] Shalom, M.; Ressnig, D.; Yang, X.; Clavel, G.; Fellinger, T. P.; Antonietti, M. *J. Mater. Chem. A* **2015**, *3*, 8171–8177.

[259] Swesi, A. T.; Masud, J.; Nath, M. *Energy Environ. Sci.* **2016**, *9*, 1771–1782.

[260] Yang, W.; Rehman, S.; Chu, X.; Hou, Y.; Gao, S. *ChemNanoMat* **2015**, *1*, 376–398.

[261] Barone, V.; Casarin, M.; Forrer, D.; Pavone, M.; Sambi, M.; Vittadini, A. *J. Comput. Chem.* **2009**, *30*, 934–939.

[262] Zhou, Z.; Zhao, J.; Chen, Z.; Gao, X.; Yan, T.; Wen, B.; Schleyer, P. v. R. *J. Phys. Chem. B* **2006**, *110*, 13363–13369.

[263] Krishnan, S.; Vadapoo, R.; Riley, K. E.; Velev, J. P. *Phys. Rev. B* **2011**, *84*, 165408 (1–6).

[264] 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. *APL Mater.* **2013**, *1*, 011002.

[265] Saal, J. E.; Kirklin, S.; Aykol, M.; Meredig, B.; Wolverton, C. *JOM* **2013**, *65*, 1501–1509.

[266] Ivanciuc, O. *Reviews in Computational Chemistry* **2007**, *23*, 291–400.

[267] Handley, C. M.; Popelier, P. L. *J. Phys. Chem. A* **2010**, *114*, 3371–3383.

[268] Nørskov, J. *Prog. Surf. Sci.* **1991**, *38*, 103–144.

[269] Hammer, B.; Norskov, J. *Nature* **1995**, *376*, 238–240.

[270] Raaen, S.; Parks, R. *Solid State Commun.* **1983**, *48*, 199–202.

[271] Sabatier, F. *La Catalyse en Chime Organique*; Berauge: Paris, France, 1920.

[272] Mavrikakis, M.; Hammer, B.; Nørskov, J. K. *Phys. Rev. Lett.* **1998**, *81*, 2819–22.

[273] Strasser, P.; Koh, S.; Anniyev, T.; Greeley, J.; More, K.; Yu, C.; Liu, Z.; Kaya, S.; Nordlund, D.; Ogasawara, H. *Nat. Chem.* **2010**, *2*, 454–460.

[274] Wang, H.; Xu, S.; Tsai, C.; Li, Y.; Liu, C.; Zhao, J.; Liu, Y.; Yuan, H.; Abild-Pedersen, F.; Prinz, F. B. *Science* **2016**, *354*, 1031–1036.

[275] Brommer, K. D.; Needels, M.; Larson, B.; Joannopoulos, J. *Phys. Rev. Lett.* **1992**, *68*, 1355–1358.

[276] Wang, X.-Q. *Phys. Rev. Lett.* **1991**, *67*, 3547–3550.

[277] Michaelides, A.; Reuter, K.; Scheffler, M. *J. Vac. Sci. Technol.* **2005**, *23*, 1487–97.

[278] Schmid, M.; Reicho, A.; Stierle, A.; Costina, I.; Klikovits, J.; Kostelnik, P.; Dubay, O.; Kresse, G.; Gustafson, J.; Lundgren, E.; et al. *Phys. Rev. Lett.* **2006**, *96*, 146102.

[279] Martin, N.; Klacar, S.; Gronbeck, H.; Knudsen, J.; Schnadt, J.; Blomberg, S.; Gustafson, J.; Lundgren, E. *J. Phys. Chem. C.* **2014**, *118*, 15324–15331.

[280] Wang, Q.; Oganov, A. R.; Zhu, Q.; Zhou, X.-F. *Phys. Rev. Lett.* **2014**, *113*, 266101.

[281] Zhou, X.-F.; Oganov, A. R.; Shao, X.; Zhu, Q.; Wang, H.-T. *Phys. Rev. Lett.* **2014**, *113*, 176101.

[282] Senftle, T. P.; Duin, A. C. v. D.; Janik, M. J. *Catal. Commun.* **2014**, *52*, 72–77.

[283] Walton, K. S.; Snurr, R. Q. *J. Am. Chem. Soc.* **2007**, *129*, 8552–8556.

[284] Han, S. S.; Furukawa, H.; Yaghi, O. M.; Goddard Iii, W. A. *J. Am. Chem. Soc.* **2008**, *130*, 11580–11581.

[285] Fetisov, E. O.; Shah, M. S.; Long, J. R.; Tsapatsis, M.; Siepmann, J. I. *Chem. Commun.* **2018**, *54*, 10816–10819.

[286] Dimitrakakis, G. K.; Tylianakis, E.; Froudakis, G. E. *Nano Lett.* **2008**, *8*, 3166–3170.

[287] Heyden, A.; Duren, T.; Keil, F. J. *Chem. Eng. Sci.* **2002**, *57*, 2439 – 2448.

[288] Snurr, R. Q.; Bell, A. T.; Theodorou, D. N. *J. Phys. Chem.* **1993**, *97*, 13742–13752.

[289] Fetisov, E. O.; Shah, M. S.; Knight, C.; Tsapatsis, M.; Siepmann, J. I. *ChemPhysChem* **2017**, *19*, 512–518.

[290] Mullen, R. G.; Corcelli, S. A.; Maginn, E. J. *J. Phys. Chem. Lett.* **2018**, *9*, 5213–5218.

[291] Müller, E. A.; Rull, L. F.; Vega, L. F.; Gubbins, K. E. *J. Phys. Chem.* **1996**, *100*, 1189–1196.

[292] Potoff, J. J.; Siepmann, J. I. *AlChE J.* **2001**, *47*, 1676–1682.

[293] Chen, B.; Potoff, J. J.; Siepmann, J. I. *J. Phys. Chem. B.* **2001**, *105*, 3093–3104.

[294] Van De Walle, A.; Asta, M.; Ceder, G. *Calphad* **2002**, *26*, 539–553.

[295] Castin, N.; Messina, L.; Domain, C.; Pasianot, R. C.; Olsson, P. *Phys. Rev. B* **2017**, *95*, 214117.

[296] Poyurovskii, L. V.; Ruban, A. V.; Abrikosov, I. A.; Vekilov, Y. K.; Johansson, B. *J. Exp. Theor. Phys.* **2001**, *73*, 415–419.

[297] Wilding, N. B. *Phys. Rev. E.* **1995**, *52*, 602–611.

[298] Marini, G. W.; Texler, N. R.; Rode, B. M. *J. Phys. Chem.* **1996**, *100*, 6808–6813.

[299] Duffy, E. M.; Severance, D. L.; Jorgensen, W. L. *J. Amer. Chem. Soc.* **1992**, 7535–7542.

[300] Pendry, J. B. *Phys. Rev. Lett.* **2000**, *85*, 3966–3969.

[301] Kelly, K. L.; Coronado, E.; Zhao, L. L.; Schatz, G. C. *J. Phys. Chem. B.* **2003**, *107*, 668–677.

[302] Hughes, M. D.; Xu, Y.-J.; Jenkins, P.; McMorn, P.; Landon, P.; Enache, D. I.; Carley, A. F.; Attard, G. A.; Hutchings, G. J.; King, F.; et al. *Nature* **2005**, *437*, 1132–1135.

[303] Christopher, P.; Xin, H.; Linic, S. *Nat. Chem.* **2011**, *3*, 467–472.

[304] Jain, P. K.; Huang, X.; El-Sayed, I. H.; El-Sayed, M. A. *Acc. Chem. Res.* **2008**, *41*, 1578–1586.

[305] Chen, X.; Schluesener, H. J. *Toxicol Lett.* **2008**, *176*, 1–12.

[306] Rovida, G.; Pratesi, F.; Maglietta, M.; Ferroni, E. *J. Vac. Sci. Technol.* **1972**, *9*, 796–799.

[307] Campbell, C. T. *Surf. Sci.* **1985**, *157*, 43–60.

[308] Carlisle, C.; Fujimoto, T.; Sim, W.; King, D. *Surf. Sci.* **2000**, *470*, 15–31.

[309] Schnadt, J.; Knudsen, J.; Hu, X. L.; Michaelides, A.; Vang, R. T.; Reuter, K.; Li, Z.; Lægsgaard, E.; Scheffler, M.; Besenbacher, F. *Phys. Rev. B.* **2009**, *80*, 075424.

[310] Giannozzi, P.; et al. *J. Phys. Condens. Matter* **2017**, *29*, 465901.

[311] Wexler, R. B.; Qiu, T.; Rappe, A. M. `https://github.com/rwexler/gcmc/`.

[312] Frenkel, D.; Smit, B. *Understanding molecular simulation: from algorithms to applications*; Elsevier, 2001; Vol. 1.

[313] Senftle, T. P.; Meyer, R. J.; Janik, M. J.; Van Duin, A. C. *J. Chem. Phys.* **2013**, *139*, 044109.

[314] Lachet, V.; Boutin, A.; Tavitian, B.; Fuchs, A. H. *Faraday Discuss.* **1997**, *106*, 307–323.

[315] Wagman, D. D.; Evans, W.; Parker, V.; Halow, I.; Bailey, S. M.; *Selected values of chemical thermodynamic properties. Tables for elements 35 through 53 in the standard order of arrangement.*; Tech. Rep.; National Standard Reference Data System; 1969.

[316] Michaelides, A.; Bocquet, M.-L.; Sautet, P.; Alavi, A.; King, D. *Chem. Phys. Lett.* **2003**, *367*, 344–350.

[317] Li, W.-X.; Stampfl, C.; Scheffler, M. *Phys. Rev. B.* **2003**, *67*, 045408.

[318] Reuter, K. In *Nanocatalysis*; Springer, 2007; pp 343–376.

[319] Jones, T. E.; Rocha, T. C.; Knop-Gericke, A.; Stampfl, C.; Schlögl, R.; Piccinin, S. *Phys. Chem. Chem. Phys.* **2015**, *17*, 9288–9312.

[320] Reicho, A.; Stierle, A.; Costina, I.; Dosch, H. *Surf. Sci.* **2007**, *601*, L19–L23.

[321] Brown, I. *Acta Crystallogr. B.* **1992**, *48*, 553–572.

[322] Van Duin, A. C.; Dasgupta, S.; Lorant, F.; Goddard, W. A. *J. Phys. Chem. A.* **2001**, *105*, 9396–9409.

[323] Tersoff, J. *Phys. Rev. B.* **1988**, *37*, 6991–7000.

[324] Tersoff, J. *Phys. Rev. Lett.* **1988**, *61*, 2879–2882.

[325] Yu, J.; Sinnott, S. B.; Phillpot, S. R. *Phys. Rev. B.* **2007**, *75*, 085311.

[326] Artrith, N.; Urban, A. *Comput. Mater. Sci.* **2016**, *114*, 135–150.

[327] Artrith, N.; Urban, A.; Ceder, G. *Phys. Rev. B.* **2017**, *96*, 014112.

[328] Bukhtiyarov, V. I.; Hävecker, M.; Kaichev, V. V.; Knop-Gericke, A.; Mayer, R. W.; Schlögl, R. *Phys. Rev. B.* **2003**, *67*, 235422.

[329] Kaichev, V.; Bukhtiyarov, V.; Hävecker, M.; Knop-Gercke, A.; Mayer, R.; Schlögl, R. *Kinet. Catal.* **2003**, *44*, 432–440.

[330] Rocha, T. C.; Oestereich, A.; Demidov, D. V.; Hävecker, M.; Zafeiratos, S.; Weinberg, G.; Bukhtiyarov, V. I.; Knop-Gericke, A.; Schlögl, R. *Phys. Chem. Chem. Phys.* **2012**, *14*, 4554–4564.

[331] Jones, T. E.; Wyrwich, R.; Bocklein, S.; Rocha, T. C.; Carbonio, E. A.; Knop-Gericke, A.; Schlogl, R.; Gunther, S.; Wintterlin, J.; Piccinin, S. *J. Phys. Chem. C.* **2016**, *120*, 28630–28638.

[332] Jones, T. E.; Rocha, T. C.; Knop-Gericke, A.; Stampfl, C.; Schlogl, R.; Piccinin, S. *ACS Catal.* **2015**, *5*, 5846–5850.

[333] Carbonio, E. A.; Rocha, T. C.; Klyushin, A. Y.; Píš, I.; Magnano, E.; Nappini, S.; Piccinin, S.; Knop-Gericke, A.; Schlögl, R.; Jones, T. E. *Chem. Sci.* **2018**, *9*, 990–998.

[334] Bocquet, M.-L.; Sautet, P.; Cerda, J.; Carlisle, C. I.; Webb, M. J.; King, D. A. *J. Am. Chem. Soc.* **2003**, *125*, 3119–3125.

[335] Linic, S.; Barteau, M. A. *J. Am. Chem. Soc.* **2002**, *124*, 310–317.

[336] Linic, S.; Barteau, M. A. *J. Am. Chem. Soc.* **2003**, *125*, 4034–4035.

[337] Zelinska, M.; Oryshchyn, S.; Zhak, O.; Pivan, J.-Y.; Potel, M.; Noël, H. *Acta Crystallographica Section E* **2007**, *63*, i158–i159.

[338] Kotomin, E. A.; Mastrikov, Y. A.; Heifets, E.; Maier, J. *Phys. Chem. Chem. Phys.* **2008**, *10*, 4644–4649.

[339] Chase, M. W. J. *NIST-JANAF Thermochemical Tables*; Journal of Physical and Chemical Reference Data Monographs; American Inst. of Physics: Gaithersburg, MD, 1998.

[340] Zumdahl, S. S.; DeCoste, D. J. *Chemical Principles*, 7th ed.; Brooks/Cole: 20 Davis Drive, Belmont, CA, 94002-3098, USA, 2013.

[341] Nørskov, J. K.; Bligaard, T.; Logadottir, A.; Kitchen, J. R.; Chen, J. G.; Pandelov, S.; Stimming, U. *J. Electrochem. Soc.* **2005**, *152*, J23–J26.

[342] Methfessel, M.; Paxton, A. *Phys. Rev. B.* **1989**, *40*, 3616–3621.