Computational Design of Interfacial Properties for Materials Discovery

by

Tara Maria Boland

A Dissertation Presented in Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy

Approved August 2022 by the
Graduate Supervisory Committee:

Peter A. Crozier, Co-Chair
Arunima K. Singh, Co-Chair
Peter Rez
Christopher Muhich
Pratik Dholabhai

ARIZONA STATE UNIVERSITY

December 2022

ABSTRACT

Interfacial interactions between materials in complex heterostructures can dominate the material's response in many modern-day energy-related devices and processes. Considerable research has been dedicated towards addressing the profound effects of interfaces. Here, first-principles-based quantum mechanical simulations are discussed to characterize the interfacial materials properties of two systems. First, density-functional theory (DFT) calculations were performed for ceramic oxide grain boundaries in undoped and doped CeO_2. Second, the development, theoretical framework, and utilization of high-throughput, workflow-based, DFT calculations are presented to model the synthesis of two-dimensional (2D) heterostructured materials. Utilizing this workflow, predictive machine learning models were created to elucidate key interface-property relationships in 2D heterostructured materials.

The DFT simulations reveal that the $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary was energetically more stable than the $\Sigma 3$ (121)/[$\bar{1}$01] grain boundary due to the larger atomic coherency in the $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary plane. The alkaline-earth metal-doped grain boundary energies demonstrate a parabolic dependence on the size of the solutes, interfacial strain, and packing density of the grain boundary. The grain boundary energies were stabilized upon Ca, Sr, and Ba doping whereas Be and Mg render them energetically unstable. The electronic density of states reveals that no defect states were present in/above the band gap. The thermodynamic trapping of oxygen vacancies in the near grain boundary region was not significantly impacted by the presence of Ca-solute ions. However, the migration energy barriers within the grain boundary core were dramatically reduced with high local Ca-solute concentrations, around 0.3 eV-0.5 eV.

Chapter 5 and Chapter 6 discusses the development of the open-source, high-

throughput computational "synthesis" based workflow package *Hetero2d* and the application of *Hetero2d* using 52 Janus 2D materials and 19 metallic, cubic phase, elemental substrates. The 438 Janus 2D-substrate pairs were analyzed by identifying substrate surfaces that stabilize metastable Janus 2D materials, characterizing their effects on the post-adsorbed 2D materials, and identifying the bonding between the 2D material and substrate. Machine learning models were applied to predict the binding energy, z-separation, and charge transfer of the Janus 2D-substrate pairs providing insight into the critical properties which factor into these properties.

This is dedicated to

A childhood dream that became a reality.

All the mentors who inspired and believed in me when I did not believe in myself.

Bradley, the best thing that I ever happened to stumble upon rock climbing.

My best friend, Alex, who has supported me through this journey with as much support as I needed.

ACKNOWLEDGMENTS

First and foremost, I extend my deepest thanks toward my doctoral advisors, Prof. Peter Crozier, Prof. Arunima Singh, and Prof. Peter Rez for their continued support, encouragement, and advisement during my Ph.D. While having three advisors has been challenging at times, the knowledge I have obtained by participating in both an experimental and computational group has been invaluable towards my career. I greatly appreciate the opportunity Prof. Peter Crozier offered me to become an embedded theorist in his experimental group as I have obtained skills I likely would not have obtained by joining other groups. I would also like to acknowledge Prof. Arunima Singh for becoming my co-advisor at the start of my 3^{rd} year. Not only has it been a continuous joy to learn from and work alongside her and the growing research group, but I wouldn't be completing my degree with confidence without her. As I look towards the future, I am exited to continue to be a part of the work I started at ASU, watching it grow, and I look forward to future projects together! I would like to acknowledge Prof. Peter Rez for his guidance during the early years of my Ph.D. I will always value his support, assistance, and many off-topic conversations that took place during our meetings. I wouldn't have been able to achieve my dreams without each of you looking after me these past 6 years.

I want to acknowledge my committee members for providing critical feedback and instructive discussions improving my research. Additionally, I would like to acknowledge Akash Patel for his dedicated work maintaining our database, website, and API. I want to thank the past and current members of both research groups for insightful conversations/assistance with different aspects of my research, namely, explaining all the technical experimental terms. I also would like to acknowledge the financial support from the National Science Foundation (grant number DMR-

1308085, DMR-1840841, and DMR-1906030) and the U.S. Department of Energy (grant number DE-SC0004954). I gratefully acknowledge the Research Computing facilities and support staff at ASU, the Extreme Science and Engineering Discovery Environment (grant number TG-DMR150006), and the National Energy Research Scientific Computing Center (contract No. DE-AC02-05CH11231) for providing HPC resources that have contributed to the research results reported here.

My acknowledgements would not be complete without mentioning the numerous mentors that shaped the person I grew up to be. I owe each of you a debt of gratitude that can only be repaid by living up to the potential you saw in me when I did not. The first of which would be my high school mathematics and physics teachers, Mrs. Cindy Iwen and Mrs. Judith Krueger, for instilling in me a love of science, a work ethic that enabled my academic success, and believing in my potential from such a young age. I would not have dared to dream of my potential without your confidence, blunt encouragement, and support during high school. Dr. Julie Abrahamson, who believed in my potential enough to nominated me for a scholarship. My undergraduate physics mentor, Dr. Yen Lee Loh, for introducing me to the field of computational science, teaching me how to code, and supporting me. Finally, Patrice Giese, my McNair Scholarship advisor and friend, for seeing my potential during my undergraduate studies enabling me to focus on my studies.

TABLE OF CONTENTS

CHAPTER Page

LIST OF TABLES

LIST OF FIGURES

Figure Page

Figure Page

Figure Page

Chapter 1

INTRODUCTION

1.1 Motivation

Materials and their interfaces have played a pivotal role in technological advancements improving the quality of life for society throughout the ages. During the Industrial Revolution society was transformed from a predominately agrarian and handicraft economy to one based on machines. Advancements in technology and materials brought about new machines and power sources standardizing the incorporation and adaptation of energy reliant devices and processes into society resulting in an ever-increasing world-wide demand for energy. Utilization of energy has influenced society in innumerable ways; for instance, over the last hundred years there has been enormous economic growth (Roser 2013; Gaye 2008), large reductions in severe poverty (Roser and Ortiz-Ospina 2013), and increased life expectancy (Roser, Ortiz-Ospina, and Ritchie 2013).

Figure 1.1 shows a breakdown of the world's consumption of energy from 1965-2020 by region. The regions with the highest energy demands originate from developed regions where the gross domestic produce and quality of life are both high. To improve the quality of life and increase the prosperity of industrialized and emerging economies, access to energy is essential. Many reports have shown that increased consumption of energy for industrialized countries (e.g. China and India) is strongly correlated with the quality of human life (Liu, Chen, and Yin 2016; Gaye 2008; Mazur 2011). As industrialized countries and emerging economies increase their standard of living

towards that of developed countries, the additional world-wide demand for energy is expected to increase. The US Energy Information Administration predicts the increase in energy consumption from these developing countries to be 47 % over the next 30 years (Gordon and Weber 2022; *Electricity Market Report–July* 2021).

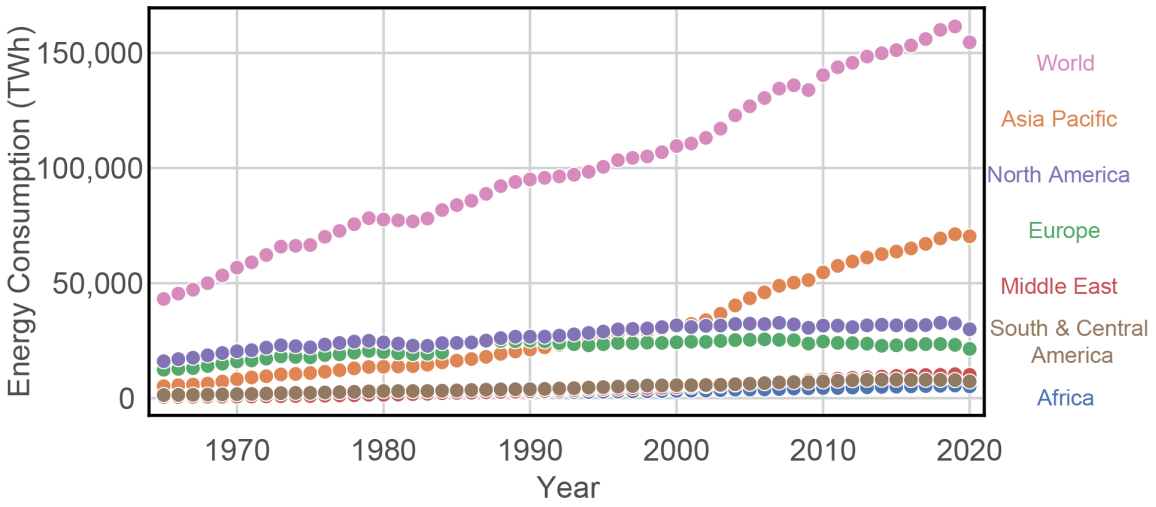

Figure 1.1. World-wide energy consumption by region from 1965-2020 in TWh. Data from Smil (2017) accessed via *Primary energy consumption by world region* (2022).

The primary energy source for developed and emerging economies are fossil fuels (*Electricity Market Report–July* 2021). While fossil fuels are a reliable and stable energy resource, they are finite resources that emit CO_2 when consumed, have been linked to anthropogenic climate change (Gaye 2008; *Electricity Market Report–July* 2021), and result in wide-spread environmental destruction during extraction (*Oil spills* 2020; *Oil Spills* 2017; Uhlmann 2020). Other sources of energy–such as renewable energy technologies–offer alternative methods to generate energy. These technologies can help curb the projected energy demand increase while also limiting the negative environmental factors that surround fossil fuel consumption.

In principle, renewable energy technologies can be used to provide continuous

energy for billions of years from a wide variety of sources. For example, Figure 1.2a shows the current global use of energy resources with the primary renewable energy sources mostly originating from traditional biomass, hydro-power, nuclear, wind, solar, and bio-fuels, listed in order of decreasing TWh energy consumption (*Global primary energy consumption by source* 2022; Smil 2017). Additionally, Figure 1.2b illustrates a renewable energy wind farm implemented in the United States. However, these technologies have some major limitations from energy production intermittency, supplemental energy storage technology requirements, low energy density, and poor efficiency compared to fossil fuels. As such, these areas have been the subject of considerable fundamental research to develop engineering strategies and new technologies for device applications (Wachsman and Lee 2011; Schmitt et al. 2020; Paul et al. 2017; Ciszewski 2012; Bondevik 2019).

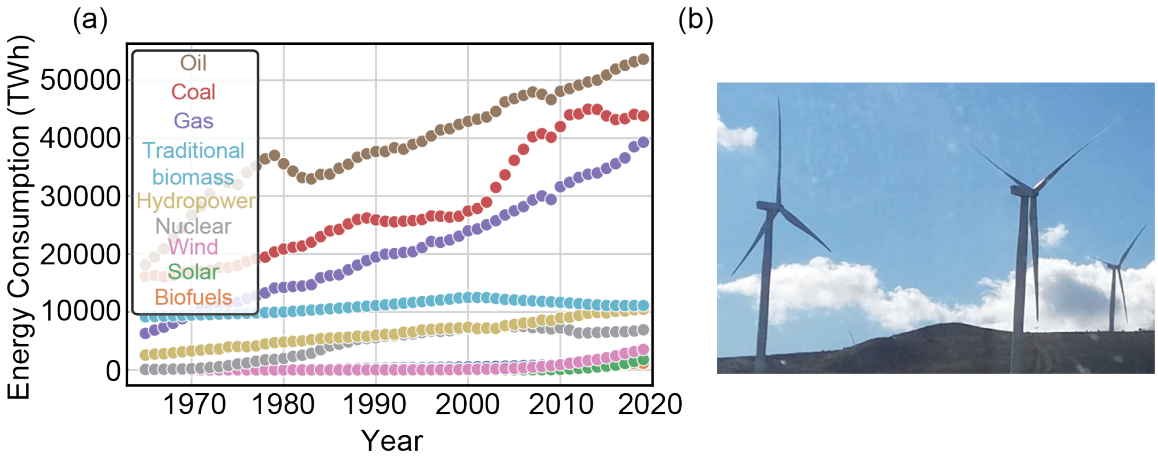

Figure 1.2. (a) The global primary energy consumption (in TWh) by source from 1965-2019. Data from Smil (2017) accessed via *Global primary energy consumption by source* (2022). (b) Renewable energy wind farm currently implemented in the United States.

Fundamental research has had substantial success in providing guidance towards engineering strategies that enable the properties of materials to be tuned for various

applications. For example, control over the optical and electronic properties of semiconductors has resulted in the development of multi-junction photovoltaic solar cells capable of adsorbing light over a significant portion of the solar spectrum (Dimroth and Kurtz 2007; Gao et al. 2022) thereby increasing the energy harvesting capability of the solar cell. Additionally, by combining concentrated solar with multi-junction solar cells, these devices have demonstrated efficiencies of 37.4% (Yamaguchi et al. 2005). In silicon photo-detectors and photo field effect transistors, precise and careful handling of the Si-SiO_2 interface lead to improvements in the UV-sensitivity for devices (Muench, Gessert, and Koeniger 1976; Ouchi et al. 1979). For polycrystalline metals, segregation of various impurities to the grain boundaries was found to cause wide-spread changes in the mechanical properties of the material from embrittlement, increased ductility, to improved mechanical strength (Braithwaite and Rez 2005; Schweinfest, Paxton, and Finnis 2004).

Despite the substantial progress in designing material properties for device applications, the interfacial properties of many materials systems often decrease the performance and inefficiency of the material. For example, one significant challenge in designing multi-junction solar cells is attaining a defect free heterointerface; typically, the interfacial regions have defects that act as recombination sites. In ion-conducting solid oxide electrolytes, the grain boundaries are often orders of magnitude more resistive than the bulk, leading to a significant decrease in total ionic conductivity for potential device applications. As such, the study, optimization, and control over the interfacial regions in complex systems (e.g. heterointerfaces and solid-solid interfaces) is one area of particular scientific interest and importance not only for renewable energy technologies, but also most modern technological applications. Polycrystalline electroceramics for use in solid oxide devices and two-dimensional (2D) heterostruc-

tured materials–that find many applications in renewable energy applications as well as many other technologically relevant fields–are two such systems which lack a fundamental understanding of materials properties that govern the interfacial interactions, shown in Figure 1.3.

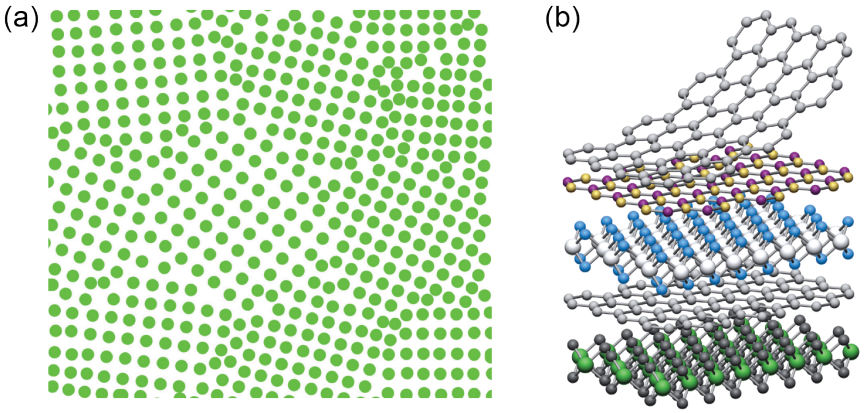

Figure 1.3. Illustrations of various types of interfaces: (a) a polycrystalline grain boundary (adapted from Pleshakov (2008)) and (b) 2D-2D heterostructured interfaces. Adapted from Geim and Grigorieva (2013) with permission from Springer Nature.

The present work was motivated by an overarching goal of characterizing the interfacial materials response to variations in critical materials properties–such as changes in the local atomic structure, composition, or energetic stability–that impact potential renewable energy applications. The following sections highlight the significant progress made towards understanding interfacial properties, the factors that modulate them, and the interdependent interface-function relationships characterizing these infinitesimal regions. These relationships can serve as the basis for future studies or aide in the development of rational design strategies and optimization of the interfacial properties in material systems with complex interfaces.

This work will focus on two interfacial system: a) grain boundaries in ceria which find applications in solid oxide fuel and electrolysis cells, and b) 2D-substrate

heterostructured material because of their considerable relevance and myriad of potential 2D-based device applications in current and future energy technologies. However, both interfacial systems presented in this work each belong to a larger class of materials systems with a diverse range of potential applications beyond renewable energy technologies (Yuan et al. 2016; Trovarelli 2002), so the computational methodology and results presented here may be generally applicable to other related systems. Nevertheless, the ubiquitous presence of interfaces coupled with the impact these regions have on the overall properties of materials systems for renewable energy applications, as well as many other modern technologies motivated this study. A brief review of interfacial systems–their expansive diversity, and challenges accompanying the study of interfaces–is discussed below, followed by a more extensive review of each system, their technological relevance, properties, and key challenges. An emphasis is placed on the interfacial interactions of grain boundaries in ceria and heterostructures formed between 2D materials and substrate surfaces.

1.2 Interfacial Systems

Complex interfaces (boundaries between solid-gas, solid-liquid, or solid-solid phases) often interact in non-trivial ways, dominating the materials response, potentially degrading the desired materials property. Interfaces, broadly defined as the common boundary between two sections of matter in space (Britannica 2022), are ubiquitous in modern devices. For some material systems, the properties of the interfacial regions can dominate the materials response leading to decreased efficiency and performance. Compounding this problem, many of these systems lack the fundamental knowledge

needed to modulate and control the interfacial properties making progress towards realization and utilization slow.

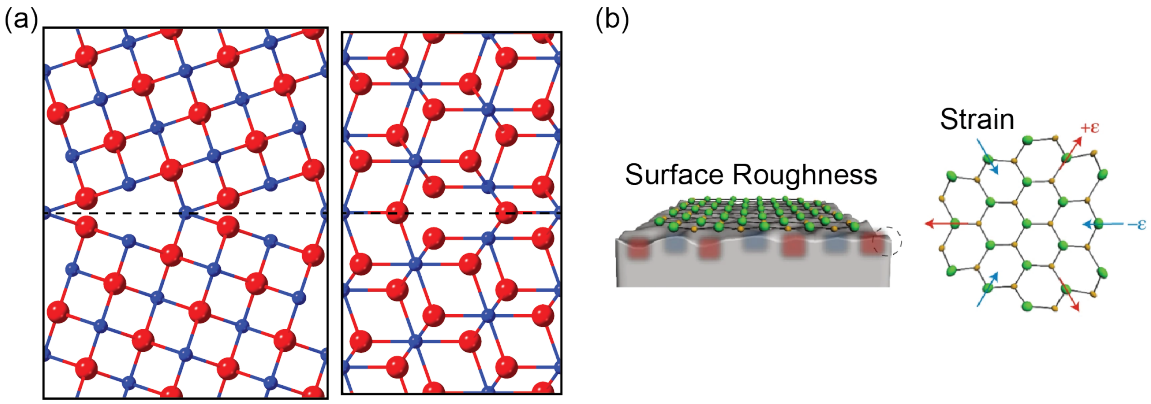

Figure 1.4. (a) Model grain boundary structure models for two grain boundary planes. (b) Illustration of surface roughness induced strain in a supported 2D material. Adapted from Rhodes et al. (2019) with permissions from Springer Nature.

The study of interfacial systems, their associated properties and impact on the overall material properties, is a diverse topic encompassing macroscopic to quantum mechanical length scales spanning numerous research fields. As mentioned previously, interfaces play a vital role in renewable energy technologies as well as many other technologies, as such, understanding how to manipulate their properties is of critical importance. A wide variety of techniques and theories can be employed to elucidate and understand the key properties that control interfacial systems and enable the manipulation of their properties. However, the study of interfacial properties is not always a straight-forward task and presents additional challenges due to the inconsistent nature of interface formation and defects. In polycrystalline materials, the solid-solid interfaces can form along any number of crystallographic planes, with differing grain sizes, and possess dramatically different local atomic structure and composition with markedly different properties compared with the adjacent grains. The growth and post-adsorbed properties of 2D materials is strongly influenced by

the growth conditions as well as the properties and defects present in the substrate which can ultimately impact the adsorbed 2D material properties.

The underlying chemistry and diverse properties that underpin interfacial regions and their associated materials functionality–coupled with multi-length scale factors that may play an important role in their overall properties–make systematic studies of these minuscule regions challenging, if not impossible in some instances. It is often necessary to utilize multiple techniques to gain fundamental insights into important factors that control interfacial properties. Utilizing multi-scale, correlative and complimentary techniques (William J Bowman et al. 2017; Bowman, Darbal, and Crozier 2020), combined computational and experimental techniques (Bondevik, Bjørheima, and Norbya 2020; Bondevik, Kuwabara, and Løvvik 2019; Pratik P Dholabhai et al. 2015), or studies which systematically change various properties and assess changes in the interfacial properties–either by computational (Aidhy, Y. Zhang, and Weber 2013; Aidhy 2016; Yuan et al. 2016) or experimental means–have had success in decoupling the intertwined interfacial properties to gain fundamental insights into the governing factors that control interfacial properties.

1.3 Ceria: Recent Progress and Outstanding Challenges

Doped polycrystalline electroceramic oxides are an important class of materials in which point defects in the bulk and grain boundaries play a key role in regulating the mechanical, optical, thermal, magnetic, catalytic and charge transport properties (Sutton and Balluffi 1995; Orlovskaya and Browning 2004; W. Lee et al. 2012; Ye et al. 2014; William J. Bowman et al. 2015; Lin et al. 2015; B. Feng et al. 2017; Nafsin et al. 2017; Moradpoor and Ebrahimi 2016). The transport and ion exchange

functionalities of electroceramics make them suitable for many technological applications including catalysts (Trovarelli 2002; Paier, Penschke, and Sauer 2013), solid electrolytes and electrodes (Schmitt et al. 2020; Trovarelli 2002; Paier, Penschke, and Sauer 2013; Singhal and Kendall 2003; Wachsman and Lee 2011; Wachsman, Marlowe, and Lee 2012), gas separation membranes (Sunarso et al. 2008; L. Zhang et al. 2012), gas sensing systems (Haile, West, and Campbell 1998), information technologies (Schweiger et al. 2017; Sunarso et al. 2008; L. Zhang et al. 2012; Haile, West, and Campbell 1998), and memristors (Schweiger et al. 2017)–among others.

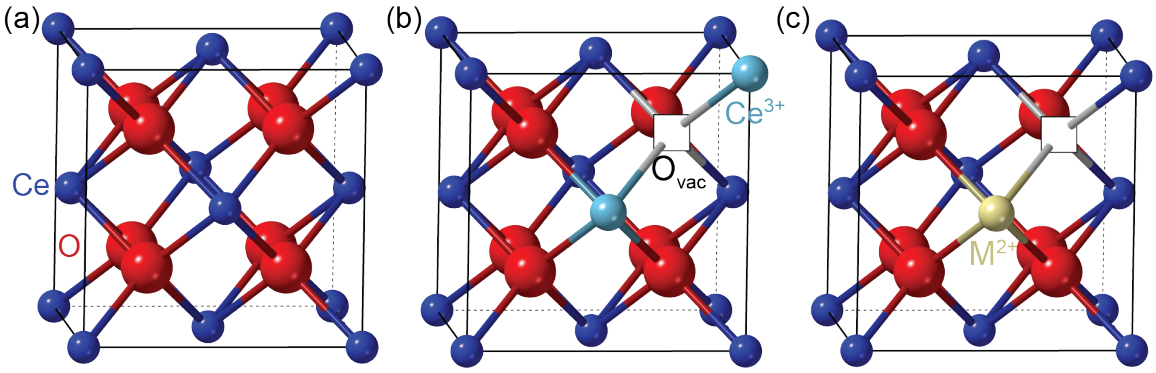

Figure 1.5. Ceria structure models illustrating the (a) pristine, (b) oxygen deficient, and (c) substitutional aliovalent cation doped crystal structure.

Most of the aforementioned technological applications rely heavily on an appreciable number of charge carriers, low electronic conductivity, and a high ionic conductivity to operate. Polycrystalline ceria–with a fluorite crystal structure (Trovarelli 2002)–and other technologically relevant oxides like Yttria-Stabilized Zirconia (YSZ) (An et al. 2013), have proven to be excellent ion-transport and ion-exchange materials when doped with a high-concentration of aliovalent cations in order to introduce charge compensating point defects. Doped ceria has gained much interest recently due to the high ionic conductivities that are achievable at low and intermediate temperatures,

thereby enabling less expensive interconnects to be used for device applications (Kwak and Jung 2016; Lee and Wachsman 2014).

In ceria, the charge carriers are positively charged oxygen vacancies (shown in Figure 1.5) which typically occupy the negatively charged lattice oxygen sites, and are transported through the electrolyte via a thermally activated vacancy hopping mechanism. Oxygen vacancies can form through various methods such as reducing conditions, elevated temperatures, and/or through doping with charge compensating substitutional aliovalent cations (Sm^{+3}, Gd^{+3}, Ca^{+2}, etc.). Aliovalent doping is an attractive avenue offering a level of extrinsic control over the material's properties in addition to providing a means to manipulate the oxygen migration energies and to regulate the concentration of mobile electrons and holes. As such, extensive work has been done to identify optimal solutes (Andersson et al. 2006; Kwak and Jung 2016; Dholabhai, Adams, et al. 2011) thereby increasing and optimizing the bulk ionic conductivity of ceria (Pratik P. Dholabhai et al. 2010a, 2010b; Jiang et al. 2005; Balducci et al. 1997; Dholabhai, Anwar, et al. 2011; Yuan, Zhang, and Weber 2015; Dholabhai and Adams 2012; Pratik P. Dholabhai et al. 2012).

While the mobile oxygen vacancies in doped polycrystalline ceria migrate relatively unimpeded through the grains, as the oxygen vacancies approach the grain boundaries, a significant reduction in the ionic conductivity is observed. For ceria and its many derivatives, optimizing the grain boundary ionic conductivity has been a persistent challenge in utilizing these materials, as these regions often control the total ionic conductivity dominating the materials response. The differing ionic conductivity has been attributed to the formation of secondary phases, variations in the so-called grain boundary character (Ye et al. 2014), and/or space charge effects which are thought to block oxygen transport across grain boundaries (Franceschetti 1981; Maier 1995;

Guo and Maier 2001; Lee and Kim 2001; Guo, Sigle, and Maier 2003; Tuller 2006; H. L. Tuller and S. R. Bishop 2010).

Considering that the ionic conductivity of the grain boundaries can be between 10^2-10^8 orders of magnitude more resistive than the grains (William J Bowman et al. 2017; Sayle, Parker, and Sayle 2006; Avila-Paredes et al. 2009; William J. Bowman et al. 2015; Bowman, Darbal, and Crozier 2020; Guo, Sigle, and Maier 2003), understanding the mechanisms which lead to increased ionic conductivity and developing strategies to increase the ionic conductivity is of critical importance to utilize these materials. However, investigating the interfacial properties can pose significant challenges. The inherent variation in grain boundary character as well as the challenges faced using either experimental or computational methods to investigate these large configuration space makes comprehensive investigations time consuming and sometimes infeasible. In the following sections, a brief review of grain boundary formalism is discussed, followed by experimental and computational methods that have made significant progress in the communities understanding of grain boundary properties.

1.3.1 Grain Boundary Formalism

A grain boundary is an extended, planar defect possessing short range order that separates regions between two grains (or crystallites), in a polycrystalline material which have different crystal orientations (Sutton and Balluffi 1995) where some atoms are not exactly aligned with either grain. These translations regions typically possess substantially different properties compared to their grain counterparts (Uesugi and Higashi 2011; Braithwaite and Rez 2005; Frechero et al. 2015; Cantwell et al. 2014; Genreith-Schriever et al. 2020) arising from the mismatch and less efficient packing

along the grain boundary region. The interfacial region composing the grain boundary can be atomically abrupt, possessing an average width between 2-5 atomic layers; they may also be more diffuse, stretching across an appreciable number of interplanar spacings (Randle 1993; Lejček 2010; Hagedorn 1973).

To describe a general interface between any two grains unambigously requires 5 independent parameters–macroscopic degrees of freedom–that describe the crystallographic relationship between the two neighbouring grains. These *macroscopic* parameters are generally referred to as the grain boundary character or geometry. These five parameters accomplish two tasks: (1) three parameters (given by the rotation axis $\boldsymbol{o}$ and angle θ) specify the mutual misorientation of the grains and (2) the last two parameters (normal to the grain boundary plane $\boldsymbol{n}$) describe the average crystallography of the boundary plane (Sutton and Balluffi 1995). These five parameters are determined *relative to one* of the grains.

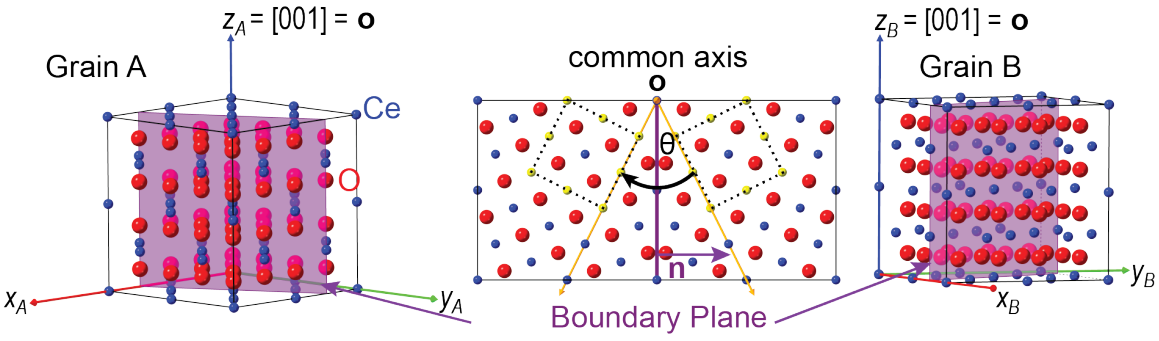

Figure 1.6. The five macroscopic degrees of freedom used to unambiguously define the grain boundary plane, mathematically given by $\theta^{\circ}[h_o k_o l_o], (h_{nA} k_{nA} l_{nA})$. The chosen coordinate system of grain A (x_A, y_A, z_A) and grain B (x_B, y_B, z_B) chosen such that the axes are parallel to crystal directions of grain A and B. The common (identical) axis in both A and B is the z-axis generally defined as $\boldsymbol{o}$= $[h_o k_o l_o]$. When one grain is rotated about $\boldsymbol{o}$ by the angle of rotation (misorientation) θ, grain A and B occupy identical positions. The grain boundary normal $\boldsymbol{n}$=$(h_{nA} k_{nA} l_{nA})$ represents the orientation of the grain boundary plane.

Figure 1.6 shows two grains (labeled Grain A and Grain B) where the shaded region of each grain illustrates the grain boundary plane. To define a grain boundary geometry, first a coordinate system must be defined for each grain where x, y, z are chosen along parallel crystallographic directions $[hkl]$ of each respective grain; labeled x_A, y_A, z_A and x_B, y_B, z_B for grain A and B, respectively. Then a common axis $\boldsymbol{o}$ is selected which, when rotated by θ, would result in a perfect matching of the lattice sites of the two grains. The normal to the grain boundary plane is then chosen for one grain denoted by $\boldsymbol{n}$=$(h_{nA}k_{nA}l_{nA})$. Combined these parameters described any general grain boundary given formally as $\theta^{\circ}[h_ok_ol_o],(h_{nA}k_{nA}l_{nA})$.

For special symmetric grain boundaries the notation can be simplified using *interface-plane* notation–used in literature and adopted in this work–combined with Σ notation given by $\Sigma(Sigma)(h_nk_nl_n)/[h_ok_ol_o]$. Σ notation describes special coincidence site lattice (CSL) grain boundaries (Cheng, Luo, and Yang 2018; Homer, Patala, and Priedeman 2015). CSL theory describes special orientations where the lattice sites of two overlapping lattices coincidence. The number following Σ relates the number of coincident lattice sites in the unit cell of the CSL to the number of lattice points in the generating lattice. An analogous interpretation of Σ is that this value relates the volume of the generating lattice to the volume of the CSL cell providing an intuitive method to understand the relationship between the original lattice and the lattice which generates the grain boundary supercell–a larger Σ value indicates a larger grain boundary lattice. In Figure 1.6 this is illustrated by outlining the generating crystal lattice with yellow atoms along the black dashed line where the CSL cell is five times that of the generating lattice (one half of the grain boundary supercell (Sutton and Balluffi 1995; Randle 1993). In essence, Σ provides an idea of how large the CSL is

and, thus, the periodic supercell one would need to create a grain boundary structure model (Talaei, Nouri, and Ziaei-Rad 2019) as discussed in later Chapter 3.

In addition to the five macroscopic degrees of freedom to consider, there are four additional *microscopic* degrees of freedom. These parameters are used to specify the manner in which the two grains are joined together–three rigid body translations that specify how the oriented grains are joined together in the x, y, z plane as well as one parameter specifying the termination of the grain boundary plane, if the system has inequivalent surface planes. These considerations are of particular importance for oxide systems due to the inequivalent surface planes and preference for anions to bond to cations.

1.3.2 Insights into Grain Boundary Properties

A wide variety of experimental (Avila-Paredes and Kim 2006; Shibata et al. 2004; Dickey, Fan, and Pennycook 2001) and computational (Aidhy, Y. Zhang, and Weber 2013; Aidhy 2016; Yang and Chen 2015; Y.-X. Feng et al. 2015; B. Feng et al. 2012) techniques have been employed to study the properties of grain boundaries in ceria and similar systems; these range from macroscopic techniques like impedance spectroscopy (Avila-Paredes et al. 2009; Sato 2015) and phase field modeling (Mebane and De Souza 2015) to nanoscale techniques like atom-probe tomography (Gregori, Merkle, and Maier 2017; X. Xu et al. 2020) and density functional theory (DFT) (Dawson, Chen, and Tanaka 2015; Yuan et al. 2016). The use of both macroscopic and nanoscale techniques have been essential to study the diverse properties of grain boundaries in polycrystalline materials. Macroscopic techniques offer statistically meaningful sample sizes capturing the ensemble average and providing a means to obtain an

overall performance metric for the material. Nanoscale characterization offers a means to characterize the markedly different chemical, compositional, structural features and properties of grain boundaries to correlate nanoscale variations with macroscopic observations.

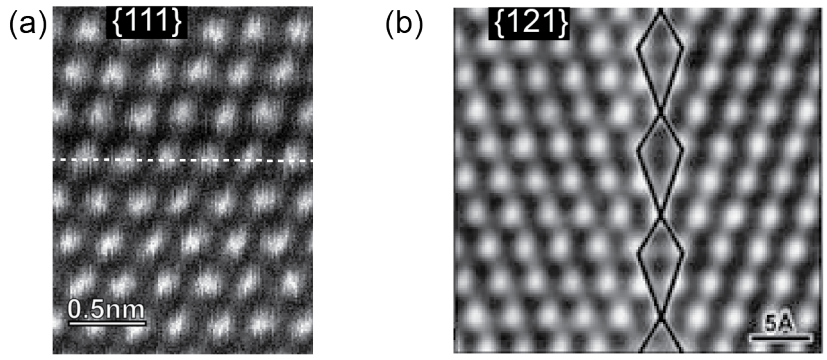

Figure 1.7. (a) A high-angle annular dark field scanning transmission electron microscopy image of the CeO_2 $\Sigma 3$ (111)/[$\bar{1}$01] and (b) high resolution transmission electron microscopy image of YSZ $\Sigma 3$ (121)/[$\bar{1}$01] grain boundary shown in grayscale. Figure copyright and citation reference from (a) B. Feng et al. (2012) adapted with permission from AIP Publishing and (b) Shibata et al. (2003) adapted with permission from Taylor & Francis Ltd.

Nanoscale characterization of ceria grain boundaries has answered many previously unknown questions regarding the local atomic structure and chemical composition of grain boundaries that were not well understood (Y. et al. 2002; An et al. 2013). As discussed previously, one possible theory attributed the decreased ionic conductivity of the grain boundary to the formation of secondary phases (Guo et al. 2013; Guo, Sigle, and Maier 2003) where subsequent high-purity polycrystalline (Guo, Sigle, and Maier 2003; Vanpoucke et al. 2014; Diercks et al. 2016) samples showed little improvement in the grain boundary ionic conductivity. Nanoscale experimental and computational characterization of model bi-crystal samples (Ye et al. 2014; Bin Feng et al. 2016; Dickey, Fan, and Pennycook 2001; Shibata et al. 2003; Shibata et al. 2004; Shibata

et al. 2002; Hojo et al. 2010) and polycrystalline materials (Frechero et al. 2015; William J Bowman et al. 2017; Bowman, Darbal, and Crozier 2020) have explored the relationship between the chemistry and local atomic structure of the grain boundary with changes in the grain boundary character (misorientation angle) such as the grain boundaries shown in Figure 1.7. These studies yielded significant insights into the nature and behavior of grain boundaries.

Direct nanoscale characterization and elemental mapping of grain boundaries (Y. et al. 2002; Orlovskaya and Browning 2004; Shirpour et al. 2012; An et al. 2013; Y.-H. Lee et al. 2013; William J. Bowman et al. 2015; Lin et al. 2015; Frechero et al. 2015; Dickey, Fan, and Pennycook 2001; Pratik P Dholabhai et al. 2015) found that solutes segregate strongly to the grain boundaries–in some cases far above both the nominal solute concentration and bulk solubility limit. The presence of high solute concentrations at the grain boundaries in ceria and similar systems has been confirmed using a variety of techniques such as atom-probe tomography (Lia et al. 2010; Diercks et al. 2016) and electron energy loss spectroscopy (William J Bowman et al. 2017; W. Lee et al. 2012). Some works have directly correlated this increase in solute concentration with increased grain boundary ionic conductivity (William J Bowman et al. 2017). These finding have shifted the focus of space charge theories from simple "dilute-solute" or "non-interacting" defect models (Guo, Sigle, and Maier 2003; Avila-Paredes and Kim 2006; Avila-Paredes et al. 2009; H. L. Tuller and S. R. Bishop 2010; Harry L. Tuller and Sean R. Bishop 2011; Kim 2016, 2016) to developing models which take into account defect-defect interactions (An et al. 2013; Lee, Prinz, and Cai 2013; Mebane and De Souza 2015) to address these deficiencies.

Despite the recent progress in theoretical and experimental characterization of grain boundary properties discussed above, the fundamental origin for the conductivity

increase with high solute concentrations remains elusive and has not been fully explored previously. Identifying and exploring the role of variable solute concentration effects on the atomic structure, bonding, energetic, and electronic properties of grain boundaries in a systematic manner may provide some guidance towards understanding how atomic level factors modulate grain boundary properties. The use of *ab-initio* computational methods, such as DFT, can be used to systematically study atomic level properties. As such, the exploration of solute composition, concentration effects, and the atomic level factors that modulate the energetic and transport properties of undoped and doped grain boundaries in ceria is the focus of Chapter 3 and Chapter 4.

1.4 2D-Substrate Heterostructured Materials

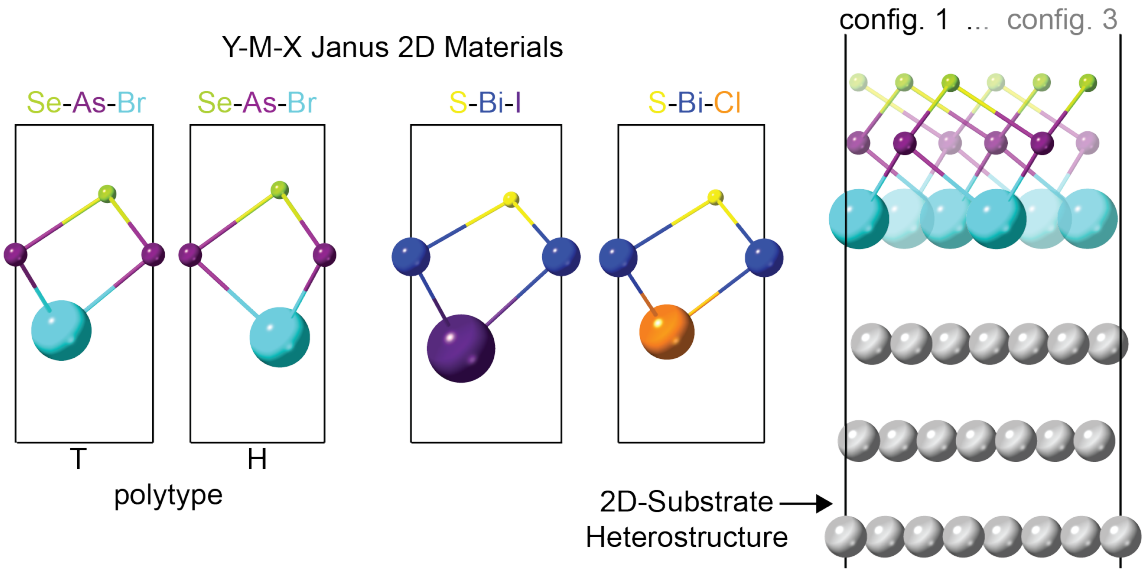

Figure 1.8. Four Y-M-X Janus 2D materials illustrating both the T and H polytype for TMDC like structures. Final structure model illustrate the T-polytype of Se-As-Br on a metallic (111) substrate surface. The differing transparencies represent the 2D material shifted relative to the substrate surface generating 3 configurations.

Two-dimensional (2D) materials are a broad class of atomically-thin materials with a wide-range of theoretically predicted chemistry's, electronic properties, and structural allotropes. These unique materials are composed of a few atomic to monolayer(s) of atoms held together by strong, in-plane covalent/ionic bonds with bond strengths of 2-8 eV per atom. In the out-of-plane direction (third dimension), between the layers, the atomically thin sheets interact through weak van der Waals (vdW) forces. These interlayer interactions are between 0.03 and 0.2 eV per atom (Deng et al. 2017), which is 1 to 2 orders of magnitude lower than the intralayer bonding. The strong in-plane intralayer bonding of a 2D crystal provides stability and a degree of chemical inertness, while the weak vdW interactions in the third dimension gives rise to novel properties originating from one-dimensional quantization. These one-dimensional quantum confinement effects (number of dimensions with quantization) give rise to the dramatically different properties 2D materials possess when compared to their bulk counterparts, layered or otherwise.

The excitement surrounding the discovery and successful isolation of single-layer graphene (Novoselov et al. 2004) sparked considerable interest in 2D materials; both experimental and theoretical research aimed to discover and synthesize (Qin et al. 2022; Y.-H. Lee et al. 2013; Kong et al. 2013) new low-dimensional materials for a myriad of promising novel next-generation device applications (Novoselov et al. 2016). These extraordinary materials have wide-ranging properties such as large band-gap insulators (Chaves et al. 2020), high-temperature superconductivity (Saito, Nojima, and Iwasa 2016), mechanical toughness (Zhang and Cheung 2016), finite out-of-plane dipole moments (G. Zhang et al. 2021; Riis-Jensen et al. 2019), chemical activity (Singh, Mathew, et al. 2015; Paul et al. 2017), and in-plane piezoelectricity (Blonsky et al. 2015). While the potential applications cover wide-ranging fields such as nanoelec-

tronics (Kang et al. 2013; Mak et al. 2010; Chen et al. 2013; G. L. Yu et al. 2014), quantum computing (K. Kim et al. 2018; S. Xu et al. 2015), field-effect transistors (Amani et al. 2014), microwave and terahertz photonics (Z. Shi et al. 2022), and catalysis (Luo, Liu, and Wang 2016; Li et al. 2019). The diverse properties and considerable research interest in 2D materials has resulted in numerous databases (Haastrup et al. 2018; Zhou et al. 2019; Mounet et al. 2018; Sorkun et al. 2020; Mathew et al. 2016) predicting new theoretical 2D crystal structures, compositions, and their associated properties. These databases rely largely on DFT calculations or similar *ab-initio* methods to determine the properties of these hypothesized 2D crystals.

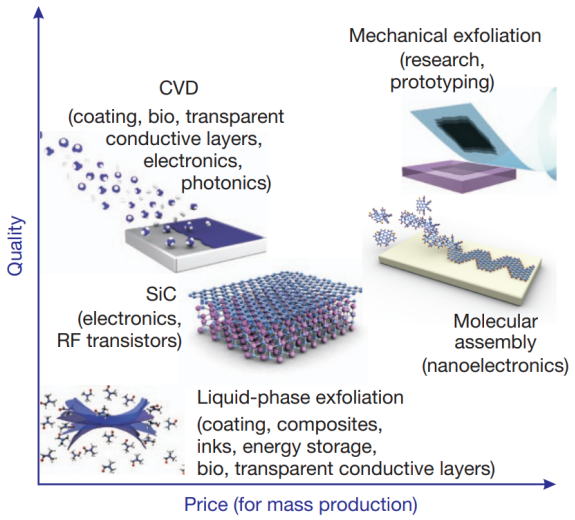

Figure 1.9. Methods for mass-production of 2D materials allowing for a wide choice in terms of size, quality and price for any particular application. Adapted from Novoselov et al. (2012) with permission from Springer Nature.

Despite the excitement surrounding this promising class of materials, surprisingly few 2D materials have been realized or utilized in device applications. Roughly 55 (Haastrup et al. 2018) of the >5,000 theoretically predicted 2D materials (Haastrup et al. 2018; Mounet et al. 2018; Ashton et al. 2017; Mathew et al. 2016; Zhou et

al. 2019) have been experimentally synthesized. The foremost technological challenges facing the realization of many as-yet-hypothesized 2D materials are: (a) reliable methods to produce low-defect materials with uniform properties at reasonably low costs, (b) the inability to stabilize the metastable 2D material (over other forms such as phases or multi-layer, nano-islands), (c) identification of suitable synthesis substrates, and (d) the unknown interaction strength between the multitude of 2D materials and substrate surfaces (Dong, Zhang, and Feng 2018; Zhang, Dong, and Ding 2021). These key challenges are discussed below; addressing the last two points is the focus of Chapter 5 and Chapter 6 which aims to identify suitable substrates and assess the post-adsorbed properties of 2D-substrate heterostructured materials.

Of the various methods available to synthesize 2D materials, substrate-assisted methods–such as chemical vapor deposition (CVD)–are one of the most technologically relevant methods; these methods result in large-area, low-defect flakes at a reasonable cost per mass (Novoselov et al. 2012; Zhang, Dong, and Ding 2021) shown in Figure 1.9. Substrate-assisted methods have the added benefit of stabilizing 2D materials with high formation energies (Singh, Mathew, et al. 2015) and the capability of synthesizing 2D materials without a vdW bonded bulk counterpart (Singh, Mathew, et al. 2015) which many as-yet-hypothesized 2D materials do not have. Yet, substrate-assisted synthesis methods face many practical challenges steaming from the expensive trial-and-error processes to synthesize 2D materials requiring significant experimental effort and intuition for choosing the substrate, precursors, and the growth conditions (the substrate temperature, growth rate, etc.). For all but the most commonly studied 2D materials (e.g. graphene, h-BN, MoS_2), this area remains in its infancy. Additionally, realization of theoretically predicted 2D materials (Qin et al. 2022) can be challenging due to their inherently metastability and unknown substrate effects (Singh, Hennig,

et al. 2015; Singh, Zhuang, and Hennig 2014) which can be non-trivial and impact the materials properties. The choice of substrate can alter the 2D materials properties as seen in 2D-MoS_2 where the mobility of carriers are reduced by more than an order of magnitude when placed on a sapphire substrate (Singh, Hennig, et al. 2015). In come cases, the successful growth of a 2D material is determine by the type of substrate; graphene and hexagonal-BN (h-BN) are stabilized by metal substrates while most transition metal dichalcogenides (TMDC) have been grown successfully on SiO_2/Si or sapphire substrates (Zhang, Dong, and Ding 2021).

Considering the immense number of theoretically predicted 2D materials, all but a subset of commonly studied 2D materials–like graphene (Kang et al. 2013; Ci et al. 2010), h-BN (Akiyama, Kawamura, and Ito 2021), or common TMDC (MoS_2, $MoSe_2$, WS_2, WSe_2) (Kong et al. 2013; Y. Yu et al. 2013; Gurarslan et al. 2014; Chen et al. 2013; Zeng et al. 2013)–have comprehensive studies (Idrees, Fawad, et al. 2020; Idrees, Din, et al. 2020; Chen et al. 2013; Deng et al. 2017) detailing the interactions of the 2D material with various substrates. Having a detailed understanding of the materials dependent properties dictating the energetic stability, nature of the 2D-substrate interaction (ionic/covalent/vdW bonding, charge transfer, 2D-substrate z-separation distance, etc.), and post-adsorbed properties of 2D materials with various substrates to overcome some of the challenges associated with the experimental realization of these unique materials. Computational methods such as DFT as well as high-throughput workflows, their associated databases–containing computed materials properties–have the capability to identify potential substrates that stabilize 2D materials and provide insight into the 2D-substrate interaction and their post-adsorbed properties. However, the combination of these two methods has not been fully addressed within the scientific community.

1.5 Goals and Objectives

The previous sections highlight the significant progress made towards understanding interfacial properties as well as the complex, interdependent interface-function relationships that characterize these regions. Despite the wealth of knowledge concerning interfacial properties, fundamental questions still remain regarding what controls these interactions of various systems, and how to design interfaces to optimize specific material properties. Considering the nanoscopic size, expansive diversity, and challenges (at both the macro- and nanoscale) associated with probing these interfacial regions, it is unsurprising that a comprehensive understanding of fundamental factors which govern these systems is a challenging task. Often, progress relies on utilizing experimental, theoretical, and computational techniques to understand and identify key interface-property relationships.

Over the past decade, the application of high-throughput computational workflows– to accelerate materials discovery and deepen the communities understanding of fundamental materials properties–has been an area of intense interest. At the heart of this so-called "Materials Genome Initiative" (MGI) is the notion that the deployment of new materials can be accelerated through the joint efforts in theory, computation, and experiment (Pablo et al. 2019). The MGI aims to share, develop, and advance the understanding of materials and their applications in a collaborative and synergistic manner. Data-science and machine learning are areas which have directly benefited from the recent interest, development, and wide-spread adoption of the infrastructure put forth by the MGI. These areas have become increasingly important within the materials science community for handling and elucidating key material-property relationships and deepening the communities understanding of materials. The work

described in subsequent chapters utilizes and further develops the tools provided by MGI–with the primary focus of understanding the interfacial properties of materials.

Ultimately, the goal of this work is to develop strategies which identify fundamental relationships and key properties that modulate the interface, and assess the impact these key properties have on overall material properties through computational approaches. Motivating research questions include:

- In ceria, how are the grain boundary properties influenced by variations in local solute concentrations and composition? Of special interest is how the presence of divalent alkaline-earth metal solutes modulate the grain boundary. Do the trends observed for divalent solutes match those found for trivalent solutes?
- In Ca-doped ceria, what is the origin of the increased grain boundary ionic conductivity with variable local solute concentrations at the grain boundary? How does the presences of solutes effect the migration energy of oxygen vacancies?
- Can high-throughput computational techniques–like those used by the Materials Genome Initiative–be utilized and extended to address fundamental questions regarding the stability and post-adsorbed properties of 2D heterostructured materials?
- What fundamental insights regarding the stability and nature of 2D-substrate heterostructure interactions can be obtained from large data sets and utilization of techniques such as machine learning? What, if any, are the fundamental material properties that govern the interfacial interactions of adsorbed 2D materials?

To this end, *ab-initio* quantum mechanical simulations characterizing fundamental materials properties are undertaken to systematically probe variations in the interfacial properties of (a) ceramic oxide grain boundaries of undoped and doped CeO_2 and (b)

heterostructures formed between any arbitrary two-dimensional material and low-index substrate surfaces. The overarching goal of Chapter 3 and Chapter 4 is to explore the compositional effects, assessing how the interfacial properties of grain boundaries in ceria are modulated. The thermodynamic stability, atomic and electronic structure of undoped and doped CeO_2 grain boundaries are examined using density-functional theory simulations, with a GGA+U functional. Two model symmetric tilt grain boundaries were created and the grain boundary core systematically doped with high local solute concentrations of alkaline-earth metals solutes. The significance of the results are discussed in Chapter 3. For Ca-doped ceria, the aim is to provide clarity regarding the origin of the increased grain boundary ionic conductivity, e.g. quantify changes in the migration energy of oxygen vacancies with high Ca concentrations at the grain boundary. The hypothesis is that high concentrations of solutes located at the grain boundary are largely responsible for modulating the ionic conductivity of polycrystalline ceramics and the results are discussed in Chapter 4. Chapter 5 discusses the development of the high-throughput computational framework to provide insights into the fundamental interfacial properties–like nature of bonding, charge doping, lattice and electronic structure distortions, and energetic stabilization–of 2D material-substrate surface interfaces. In Chapter 6, the high-throughput workflow is combined with well-known machine learning techniques to elucidate fundamental factors governing the substrate induced stabilization of Janus 2D materials on metallic substrate surfaces.

Chapter 2

RESEARCH METHODOLOGIES

Density functional theory (DFT) is a common method from which the energetic, thermodynamic, and kinetic properties of experimentally derived and hypothesized material structures can be explored and/or determined while accounting for quantum mechanical effects. DFT-based methods have had remarkable success describing the properties of materials, guiding experiments, and elucidating key material-property relationships. These methods are relatively accurate in describing the properties of a diverse range of materials–from ionically bonded to vdW bonded materials–which can be challenging to obtain with computationally less expensive methods, such as molecular dynamics. The versatility of DFT and the recent development of high-throughput computational workflows has been of great benefit to the scientific community for materials discovery. Naturally following the utilization of these techniques, large databases containing a number of computed materials properties have emerged opening new avenues to utilize long standing methods from other fields such as data-science and artificial intelligence. This section covers the computational methodologies utilized in the subsequent chapters from DFT, workflow development, data provenance, and the application of machine learning to study the properties of grain boundaries and 2D-substrate heterostructures.

2.1 Density Functional Theory

Density-functional theory (DFT) is, in principle, an exact method to solve the many-bodied Schrödinger equation describing material properties arising from the *ground state* interactions of electrons and nuclei. The time-independent, non-relativistic Schrödinger equation is given simply by Equation 2.1 where $\widehat{H}$ is the Hamiltonian operator, Ψ represents the total wave-functions (eigenstates), and E is the associated energy of each wave-function (eigenvalue) of the system.

$$\widehat{H}\Psi = E\Psi \tag{2.1}$$

The revolution of DFT and the exactness of the theory comes from the realization that instead of solving an intractable, many-bodied problem (i.e. solving Equation 2.1 with $\widehat{H}$ given by Equation 2.2), one could first separate the electronic and ionic degrees of freedom and then map the system of N interacting electrons to a unique functional of electron density. In principle, this turns out to be sufficient to solve for the exact quantum mechanical **ground state** energy of the interacting electron system which is covered in more detail in the following paragraphs.

The many-bodied Hamiltonian operator is given by Equation 2.2 where $\hbar$ is the reduced Planck constant, M_I is the mass of the I^{th} nucleus at positions $\mathbf{R}_I$ with nuclear charge eZ_I, m_e is the mass of the n^{th} electron at positions $\mathbf{r}_i$ with charge e.

$$\widehat{H} = -\sum_{\mathrm{i=1}}^{\mathrm{n}} \frac{\hbar^2}{2\mathrm{m_e}}\nabla^2_{\mathbf{r_i}} - \sum_{\mathrm{I=1}}^{\mathrm{N}} \frac{\hbar^2}{2\mathrm{M_I}}\nabla^2_{\mathbf{R_I}} + \sum_{\mathrm{i>j}} \frac{e^2}{|\mathbf{r_i} - \mathbf{r_j}|} - \sum_{\mathrm{i=1}}^{\mathrm{n}}\sum_{\mathrm{I=1}}^{\mathrm{N}} \frac{\mathrm{Z_I}e^2}{\mathbf{R_I} - \mathbf{r_i}} + \sum_{\mathrm{I>J}} \frac{\mathrm{Z_I Z_J}\, e^2}{\mathbf{R_I} - \mathbf{R_J}} \tag{2.2}$$

In order from left to right, the terms in Equation 2.2 represent the contribution to the total energy originating from: the kinetic energy (momentum) of the electrons and

nuclei, the energy of Coulomb interaction (electrostatic) between electrons, between electrons-nuclei, and between nuclei.

The first step towards simplifying Equation 2.2 is to decouple the interactions between the electrons and nuclei. The work of Born and Oppenheimer (Born and Oppenheimer 1927) laid the theoretical foundation to identify a mathematically tractable solution for Equation 2.1 and Equation 2.2. Later works by Kohn and Hohenberg determined a ridge formulation of DFT which forms the basis of modern DFT (Hohenberg and Kohn 1964).

The Born-Oppenheimer approximation simplified Equation 2.2 separating the ionic and electronic degrees of freedom by arguing the vibrational frequencies of the electrons are orders of magnitude larger than the ions. This observation lead to the belief that ions appear frozen to the electrons. Following the Born-Oppenheimer approximation, the work of Kohn and Hohenberg mapped the electron many-body ground state wave-function of a system of N interacting electrons to a unique functional of the electron density. Their work enabled the total energy functional of the many-body, interacting-electron system (with 3N degrees of freedom) to be written as a sum of the external potential and the electronic energies (with 3 degrees of freedom) shown in Equation 2.3 where $T[n[(\mathbf{r})]$ is the kinetic energy, U_{e} is the electron-electron interaction, and V_{ext} is the external potential.

$$E[n(\mathbf{r})] = \langle \Psi | \, \widehat{H} \, | \Psi \rangle = T[n(\mathbf{r})] + U_{\mathrm{e}}[n(\mathbf{r})] + \int V_{\mathrm{ext}}(\mathbf{r}) n(\mathbf{r}) d\mathbf{r} \tag{2.3}$$

Another simplification proposed by Kohn and Sham (Kohn and Sham 1965) theorized that the kinetic energy term can be represented using one-electron orbitals. This enabled the sum of one-electron orbitals to represent the electron density of a system with N interacting electrons (Lesar 2013) and to redefine the kinetic energy functional as $T_{KS}[n(\mathbf{r})] = \sum_{i=1}^{N} \langle \Psi_i | - \frac{1}{2} \nabla_i^2 \, | \Psi_i \rangle$ where Ψ_i represents the one-electron

orbital wave-function. This assumption reduced the problem from a multi-electron problem to one where you would find a set of solutions to one-electron problems.

Furthermore, they proposed that the electron-electron interaction energy ($U_{\mathrm{e}}[n(\mathbf{r})]$) can be represented by the classical electrostatic (Hartree) energy. To account for the inaccuracies associated with their assumptions a correction term, known today as the exchange-correlation functional $E_{\mathrm{xc}}[n(\mathbf{r})]$, was introduced. $E_{\mathrm{xc}}[n(\mathbf{r})]$ originates from (1) the kinetic energy difference between the interacting and non-interacting systems and (2) the non-classical electrostatic interaction energy. If the exchange-correlation functional is known, then (in principle) DFT is an exact theory. However, the exact form of the exchange-correlation functional is typically unknown, making practical DFT applications an approximate solution. The general form for the exchange-correlation functional is given by Equation 2.4.

$$E_{\mathrm{xc}}[n(\mathbf{r})] = T[n(\mathbf{r})] - T_{\mathrm{KS}}[n(\mathbf{r})] + U_{\mathrm{e}}[n(\mathbf{r})] - J[n(\mathbf{r})] \tag{2.4}$$

2.1.1 Approximations to the Exchange-Correlation Functional

There are many ways to approximate the exchange-correlation functional; one can categorize each approximation using Jacob's ladder shown in Figure 2.1 (Burke 2012). The lower rungs are more simplistic approximations starting in the Hartree (Coulomb) world, where the N-particle wave-functions are represented as a product of single-particle wave-functions that neglect non-classical interactions arising from inhomogeneities in the electron gas. The local density approximation (LDA) to the exchange-correlation functional falls into this category where the local electron energy density at any position $\mathbf{r}$ is replaced with the exchange-correlation energy associated

with that of the homogeneous electron gas of the same density (Lesar 2013; Burke 2012).

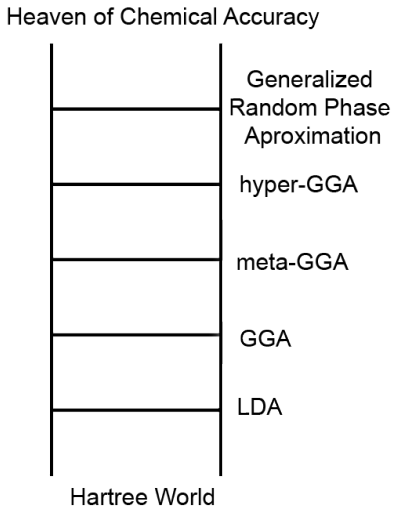

Figure 2.1. Jacob's ladder of density functional approximations to the exchange-correlation energy. Adapted from Perdew et al. (2005) with permission from AIP Publishing.

The second rung on the ladder takes the gradient corrections of $n(\mathbf{r})$ and is known as the generalized gradient approximation (GGA) (J. P. Perdew 1985). The third rung is reserved for meta-GGA functionals which include functionals such as TPSS (Tao et al. 2003)–which refers to the authors initials. The fourth rung contains hybrid functionals that mix some exact exchange with GGA such as the B3LYP (Stephens et al. 1994) or HSE0 (Krukau et al. 2006; Heyd, Scuseria, and Ernzerhof 2006; Vydrov and Scuseria 2004; Heyd, Scuseria, and Ernzerhof 2003) functionals. The final rung is fully non-local functionals, such as the random phase approximation (Dobson, Wang, and Gould 2002; Fuchs and Gonze 2002) or B2PLYP (Grimme and Neese 2007) functionals. However, these functionals are extremely computationally expensive to use.

For many materials systems, these initial approximations to the exchange-correlation functional successfully describes the ground state properties. However, for systems which posses strongly correlated electrons (e.g. Mott insulators or compounds that contain partially filled $3d$ or $4f$ states) or London dispersion interactions (e.g. systems with van der Waals interactions like two-dimensional materials) the general prescription for the exchange-correlation functional fails (Sõderlind et al. 2010). For these systems, modifications to the exchange-correlation functionals have been made to more accurately represent the physical system.

In correlated systems, approximations to the exchange-correlation functional do not completely cancel out the electronic self-interaction contained in the Hartree term. This results in a fragment of the same electron inducing an added self-interaction energy that results in excessively delocalized wave-functions. This is the origin of the tendency for DFT to over-delocalize valence electrons and over-stabilize metallic ground states, resulting in systematic errors. The Hubbard model (H. and K. 2002) was proposed as a solution for correcting inaccurate cancellations of the self-interaction energy for the existing formulations of the exchange-correlation functionals. Of the proposed corrective methods, the rotationally invariant energy functional proposed by Dudarev et al. (1998) is used within this work where the total energy can be written simply as

$$E_{\mathrm{DFT+U}} = E_{\mathrm{DFT}} + \frac{U-J}{2}(\sum_{j,l,\sigma} \rho^{\sigma}_{j,l}\rho^{\sigma}_{l,j}) \tag{2.5}$$

where ρ is the on-site density matrix of the d/f electron. In this formulation, only the difference between U (the effective on-site Coulomb interaction parameter) and J (the effective on-site exchange interaction parameter) is meaningful.

For materials with long-range or non-local interactions, standard approximations to the exchange-correlation functional cannot capture interactions originating from

fluctuations in the charge density distribution. This failure to describe the interactions originating from non-local charge density fluctuations is inherently due to the assumption that the electron density can be completely described as a functional of the local electron density or the gradient of the electron density. Since vdW interactions are purely a non-local interaction and traditional DFT assumes only local contributions to the electron correlation are needed, these interactions are neglected. As such, various schemes for approximating the contribution of the electron correlation in the exchange-correlation functional have been developed to include these long-range, non-local interactions. The van der Waals density functional (vdW-DF) based directly on the electron density is the most appealing approximation for solving the non-local contributions to the exchange-correlation functional. This method adds a correction to the conventional Kohn-Sham DFT energy, E_{DFT}, enabling the total energy to include contributions from dispersion, shown in Equation 2.6.

$$E_{\mathrm{DFT-disp}} = E_{\mathrm{DFT}} + E_{\mathrm{disp}} \tag{2.6}$$

$$E_{\mathrm{XC}} = E_{\mathrm{X}}^{\mathrm{GGA}} + E_{\mathrm{C}}^{\mathrm{LDA}} + E_{\mathrm{C}}^{\mathrm{nl}} \tag{2.7}$$

Dion et al. (2004) proposed one vdW-DF approximating E_{disp} through a non-local correlation functional given by Equation 2.7, where $E_{\mathrm{X}}^{\mathrm{GGA}}$ is the exchange energy based upon the revPBE GGA functional (Zhang and Yang 1998), $E_{\mathrm{C}}^{\mathrm{LDA}}$ is the LDA correlation energy, and $E_{\mathrm{C}}^{\mathrm{nl}}$ is the non-local energy term approximating electron correlation effects. $E_{\mathrm{C}}^{\mathrm{nl}}$ is solved using the algorithm proposed by Roman-Perez and Soler (Román-Pérez and Soler 2009), which increases the computational efficiency by transforming the double real space integral into reciprocal space. Recent work by Klimeš, Bowler, and Michaelides (2009) and Bowler and Michaelides (2011) has produced more accurate exchange functionals such as optB88-DF. The work described

in subsequent chapters makes use of these more recent advancements to describe the non-local interactions of two-dimensional heterostructures using the optB88-DF.

2.1.2 Implementation: VASP

Within this work, all DFT calculations were performed using the Vienna $Ab-Initio$ Simulation Package (VASP) (Kresse and Hafner 1993, 1994; Kresse and Furthmüller 1996a, 1996b). The DFT framework above provides an avenue from which the ground-state properties of materials can be determined using a first-principles approach. Using VASP, determination of the energetic/thermodynamic stability and kinetic properties for a diverse pool of material systems is possible. For interfacial structures–such as grain boundaries and 2D-substrate heterostructures–first principles methods enable interfacial properties to be characterized in a controlled manner, which is very challenging if not impossible to achieve via experimental methods.

In this work, the ground-state properties of various grain boundaries were used to assess the energetic/thermodynamic stability of grain boundary structures to assess how changes in local solute concentrations modulate fundamental grain boundary properties, such as structural, compositional, atomic, and electronic variations. For 2D heterostructures, the ground-state properties were used to ascertain the thermodynamic stability of various 2D materials on metallic substrates. This enables the identification of suitable substrates from which theoretically predicted 2D materials can be synthesized. Additionally, substrate effects on various 2D materials is explored to assess how the interfacial interactions impact the properties–structural, chemical, electronic, and charge transfer–of post-adsorbed 2D materials leading to actionable information to experimental synthesis and functionalization of novel 2D materials.

2.2 High-throughput Workflows

The term "high-throughput" is typically used to describe methods were automation is needed in order to complete a large number of repetitive, but well-defined tasks that are infeasible or time-consuming to complete manually. The basic premise of developing high-throughput computational workflows is to tackle these tasks and reduce the probability of introducing errors during tedious/time-consuming tasks. The development of these methods have wide-ranging benefits and uses over a various domains and skill levels from experts to novices.

A high-throughput computational workflow is the high-level name given to describe the series of steps taken to complete a main objective. These tasks could be anything associated with the completion of the task i.e. tracking file input/output, error correction, program execution, job management, analysis, data dependencies, etc. When a workflow is coupled with the predictive capabilities of DFT, the workflow's main objective may be obtaining any number of specific materials properties from computing the ground-state electronic band-structure and density of states to determining excited state properties. The easy of use, potential and scalability of these methods to accelerate materials discovery and standardize computational tasks has resulted in wide-spread adaptation and utilization with materials science. The following sections discuss the key concepts and considerations involved in the design and development of high-throughput computational workflows and data storage. A graphical illustration of the tasks and considerations is outline in Figure 2.2.

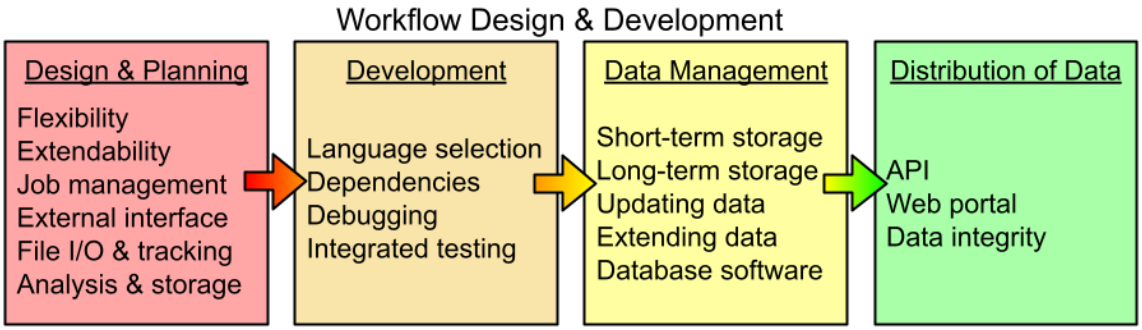

Figure 2.2. Diagram illustrating the number of steps involved in designing and developing a workflow.

2.2.1 Design and Development

The implementation of workflows often requires considerable forethought to generate an easy-to-use code that is flexible, modular, and generally compatible with other codes and across heterogeneous computational infrastructures. Additionally, the software should be capable of handling a multitude of computational tasks in a comprehensive manner that shields the user from underlying complexities; this can result in a demanding development process.

A good computational workflow is designed in a modular fashion wherein individual components are linked together and executed by a computational engine with a key characteristic being the separation between the workflow steps and its execution. The individual components linked together in a workflow should provide sufficient metadata and provenance to fully describe the generated data in a completely traceable way while the workflow itself encapsulates the methodology behind each step in the workflow. This methodology provides a flexibility framework to develop and design a wide variety of workflows that are reusable, reproducible, and easily extendable.

Due to the wide-spread popularity of workflows, a number of high-throughput

computational frameworks have been developed. These packages are capable of handling the multitude of tasks necessary for creating computational workflows such as interfacing with external software packages and computing platforms, handling computational heterogeneity, data management, and security. Packages and frameworks relevant to this work will be further discussed in Chapter 5.

Central to any software design and development is testing and validation of individual components. Testing of numerical parameters such as the calculation input and output or function and numerical validation is key to ensure the results are reasonable and to maintain FAIR (Goble et al. 2020) coding practices. Maintaining sufficient testing provides a level of future proofing as the code base and dependencies are updated.

Data provenance refers to the detailed record or history of the data and its origins. Typically this record is contained within the metadata and stored alongside related information in the database or repository. Databases are an irreplaceable component of any high-throughput computational workflow or large data set providing a means to efficiently query the generated data, direct future studies, and disseminate results by serving as the back-end to graphical websites.

2.3 Machine Learning Methods

Machine learning is a field of computer science which utilizes computer systems and efficient algorithms to "learn" for themselves from examples, i.e. large data sets. In essence, a machine learning program attempts to identify simple rules or functions that best describe a target property. Some machine learning algorithms achieve this

goal by optimizing an object function while other algorithms use a reward/punishment concept.

There are many types of machine learning algorithms and systems ranging from supervised/unsupervised to instance-based/model-based learning. These algorithms find suitable applications for wide-ranging problems in materials science. Generally, a specific algorithm is chosen based on the type of problem to be solved. In the field of materials science, some of the more popular methods are supervised learning or deep learning methods (Liu et al. 2017), as they are often easy to implement, understand, and obtain critical relational information regarding the underlying parameters describing a target property.

The utilization of machine learning for materials discovery and exploration of fundamental materials-dependant relationships requires numerous steps, such as obtaining sufficiently large data sets to train the algorithm, pre-processing the data set, building the model, and evaluation of the models performance. Additionally, this process is highly non-linear and often times requires repeating several of the steps to obtain a well-trained, generalized model. Data collection and pre-processing are the most critical steps to identify key features and obtain reasonable model performance. If the data set is not sufficiently large, consists of many highly correlated/sparse features, or does not contain predictive features, then the models performance, reliability, and generalizability will be degraded. The following sections discuss the key steps and concepts involved with generating quality machine learning models that are utilized and discussed in Chapter 6.

2.3.1 Data Collection and Pre-processing

Data collection and pre-processing is one of the most time consuming processes involved in machine learning. It requires the generation and visualization of crucial features in the data, identification of outliers, encoding the data into an appropriate format for the machine learning algorithms, and understanding data context (e.g. identifying explicit and implicit dependencies or correlations between features in the data set that impact the performance of the model). Additionally, filling in missing values that are null or removing the null values must be performed. These steps are critical to generate high quality, generalized models as most data originates from both computational simulations and experimental measurements, resulting in data sets that can be noisy and inconsistent in addition to having incomplete features.

The work discussed in Chapter 6 implements multiple methods for data collection, cleaning, and pre-processing methods. The data collection process utilizes two methods: (1) generation of *ab-initio* data using DFT, and (2) collection of data from other databases utilizing community software to apply materials based properties (features) for each data entry in the data set. Once a data set has features added to each entry it is important to compare the features and remove features which are highly correlated or are nearly homogeneous in value (most of the values in the column have the same value). The removal of highly correlated features improves the performance of the model while removing homogeneous features speeds up the training of the model.

2.3.2 Model Selection

When applying machine learning algorithms to solve a problem, the problem falls into one of two categories: (a) classification, or (b) regression. The application of regression (Liu et al. 2017; Fischer et al. 2020; G. Zhang et al. 2021; Rhone et al. 2020; S. B. Torrisi et al. 2020) (linear, random forest, or gradient boosting trees) and kernel-based (support vector machine) (Siriwardane et al. 2020; Vivanco-Benavides et al. 2022) models have been widely used in materials science to successfully predict various materials properties. These models perform well on small data sets with high-dimensional (large) feature spaces (Barnard et al. 2019; Geurts, Ernst, and Wehenkel 2006) making them particularly well suited for materials science applications.

In materials science regression algorithms are commonly used to predict an energetic quantity for a material. As such, regression algorithms such as random forest regression models are very common. Random forest models have many advantages such as (a) clear data partitioning making them more straight forward to interpret the results, (b) fast and flexible to implement for a wide range of machine learning problems, and (c) require little in the way of tuning hyper-parameters compared to other methods. Random forests are ensemble classifiers which construct multiple independent random tree models. A random tree model is a machine learning model which splits the input data into subsets on which the decision trees are trained and an aggregate model is created (Breiman 2001) leading to a more stable and generalizable results.

2.3.3 Model Training and Evaluation

Due to the easy of use and good performance on relatively small data sets, this work focuses primarily on the use of random forest regression models. Additionally, tuning the hyper-parameters of these models is easier than other methods since they have few parameters to tune. The hyper-parameters which impact the model performance the most are: (1) the number of decision trees (*n_ estimators*), and (2) the number of features considered by each tree to split a node (*max_ features*). An efficient method to tune the hyper-parameters within *scikit-learn* is to use *GridSearchCV* or *RandomizedSearchCV* which loops over specified hyper-parameter to find the optimal values for training the model.

The accuracy of random forest regression models are assessed using several statistical measures, most commonly, the coefficient of determination (R^2), mean absolute error (MAE), and root mean squared error (RMSE) as statistical measures given by Equation 2.8, Equation 2.9, and Equation 2.10, respectively. R^2 indicates the goodness-of-fit for the ML model in predicting the target values and represents the proportion of the variation in the predicted/target variable from the input (feature) variables used to predict the target value. In order to train the model K-fold cross-validation is where the training set is partitioned into K subsets and K-1 subsets train the model with the remaining subsets used to evaluate the model performance. This procedure is repeated until each subset is selected for validation (Buitinck et al. 2013).

$$R^2 = 1 - \frac{\sum_i^N (Y_{DFT}^i - Y_{predicted}^i)^2}{\sum_i^N (Y_{DFT}^i - \bar{Y})^2} \tag{2.8}$$

In Equation 2.8, the values predicted using DFT and machine learning models are denoted by Y_{DFT} and $Y_{predicted}$, respectively. $\bar{Y}$ is the average of the target values

of the test set, and N is the number of materials considered to calculate the MAE and RMSE. MAE is the absolute value of the difference between an observed value (ML predicted) of a quantity and the true value (DFT predicted). MAE provides an average magnitude of error in the ML model from the true value. RMSE provides a measure of how concentrated the data is around the line of best fit.

$$MAE = \frac{1}{N}\sum_{i}^{N} |Y^{i}_{DFT} - Y^{i}_{predicted}| \tag{2.9}$$

$$RMSE = \sqrt{\sum_{i}^{N} \frac{(Y^{i}_{DFT} - Y^{i}_{predicted})^2}{N}} \tag{2.10}$$

Chapter 3

IMPACT OF ALIOVALENT ALKALINE-EARTH METAL SOLUTES ON CERIA GRAIN BOUNDARIES: A DENSITY FUNCTIONAL THEORY STUDY

3.1 Motivation

As discussed in Section 1.3 of the Introduction, there is a need to develop a fundamental understanding of how high local solute concentrations at the grain boundary modulate the interfacial properties. In this chapter, using first-principles simulations, this work provides a fundamental understanding of the atomic-structure, energetic stability, and electronic properties of pristine (undoped) as well as aliovalent, alkaline-earth metal (AEM) doped grain boundaries in ceria. This work demonstrates that a local doping of $\sim$20% $[M]_{GB}$ (M=Be, Mg, Ca, Sr, and Ba) has a significant impact on the energetic stability of the grain boundaries. Using density-functional theory simulations with a GGA+U functional, this work examines the structure, energetic stability, and coordination of atoms at the grain boundary interface for two more frequently observed grain-boundaries in Ca-doped ceria, (Bowman 2016; William J Bowman et al. 2017) the Σ3 (111)/[$\bar{1}$01] and Σ3 (121)/[$\bar{1}$01] grain boundary.

This work finds the Σ3 (111)/[$\bar{1}$01] grain boundary is energetically more stable than the Σ3 (121)/[$\bar{1}$01] grain boundary due to the larger atomic coherency in the Σ3 (111)/[$\bar{1}$01] grain boundary plane. This work shows that a local doping with $\sim$20% $[M]_{GB}$ (M=Be, Mg, Ca, Sr, and Ba) has a significant impact on the energetic stability of the grain boundaries. As the atomic radii of the solute atom increases, the grain boundary energies display a parabolic dependence which is modulated by

the interfacial strain and the packing density of the grain boundary. In this work, a stabilization of the grain boundaries upon Ca, Sr, and Ba doping is shown whereas Be and Mg render them energetically unstable.

The element-projected and orbital-projected density of states show that no defect states are present in or above the band gap of the AEM doped ceria, which is conducive to maintaining lower electronic mobilities necessary for good ionic transport. The electronic properties, unlike the energetic properties, exhibit complex inter-dependence on the structure and chemistry of the host and the solutes. In addition, this work finds that the band gap of ceria can be modulated by up to 0.3 eV by selecting different AEM solutes at the ceria grain boundary. This work makes advances in the atomic-scale understanding of aliovalent cation doped ceria grain boundaries, serving as an anchor to future studies that can focus on understanding and improving ionic-transport.

3.2 Computational Methods

All simulations are based on DFT using the projector augmented wave method (Blöchl 1994; Kresse and Joubert 1999) as implemented in the plane-wave code VASP (Kresse and Hafner 1993, 1994; Kresse and Furthmüller 1996a, 1996b). All simulations included spin-polarization and the generalized gradient approximation (GGA) with the Perdew-Burke-Ernzerhof (PBE) (Perdew, Burke, and Ernzerhof 1996, 1997) exchange-correlation functional. In addition, the strong correlation effects of the Ce $4f$ electrons were treated within GGA using the Hubbard U correction (GGA+U) formulated by Dudarev et al. (1998). An on-site Coulomb interaction, $U_{eff} = 5$ eV, was used for Ce, as determined by Pratik P. Dholabhai et al. (2010b) as well as many others (Koettgen and Martin 2019, 2020; Pratik P. Dholabhai et al. 2010a;

Zhu et al. 2020), to provide a better fit with the experimental band gap (E_{gap}), lattice parameter (a_0), and bulk modulus (B_0) compared to traditional GGA methods. For a 2x2x2 supercell of bulk ceria, the $E_{\text{gap}}\big[\text{O}(2p) \rightarrow \text{Ce}(4f)\big]$= 2.0 eV, a_0=5.494 Å, and B_0=180.59 GPa and the corresponding experimentally measured values are $E_{\text{gap}}\big[\text{O}(2p) \rightarrow \text{Ce}(4f)\big]$=3 eV (Gerward and Olsen 1993), a_0=5.411 Å (Eyring 1979), and B_0=204-236 GPa (Nakajima, Yoshihara, and Ishigame 1994; Gerward and Olsen 1993). The chosen value of U_{eff} correctly describes the localization of the $4f$ electrons on the nearby Ce atoms–unlike traditional GGA which results in delocalized electrons on all cerium ions in the lattice.

For ceria, standard DFT fails to describe the insulating behavior and the use of the Hubbard (U) parameter that prioritizes electron localization on the nearby Ce $4f$ states is essential to correct for this. This electron localization is well established in the experimental literature that characterizes so many processes in ceria based systems (Trovarelli 2002). The DFT treatment of ceria is well established within the GGA+U method and typically overestimates the lattice parameters with errors of 1%-2% (He et al. 2014). The computed lattice parameter value is overestimated by 1.53%, well within the expected accuracy of the DFT method as well as recently published literature (Koettgen and Martin 2020, 2019; Wu, Vegge, and Hansen 2019).

Band gap underestimation within DFT is also a well-established trend (J. Perdew 1985; Jie et al. 2019; Dittmer et al. 2019) that can produce discrepancies upwards of 1-2 eV (Dittmer et al. 2019) or a 40%-50% difference between calculations and experimental measurements (Brothers et al. 2008), dependent upon the class of the material. However, this is a systematic error present in all calculations due to the band gap underestimation error (Tolba et al. 2018; J. Perdew 1985; Dittmer et al. 2019) and the choice of the exchange-correlation functional (Chan and Ceder 2010; Heyd

et al. 2005; Setyawan et al. 2011). Due to this systematic error, although absolute band gap energies might not agree with experimental measurement, differences from structural changes follow trends expected from experimental measurements. The values of U, the lattice parameter and band gap are shown in Appendix Table A.1 for comparison with other GGA+U calculations (Koettgen and Martin 2020, 2019; Zhu et al. 2020; Pratik P. Dholabhai et al. 2010a; Wu, Vegge, and Hansen 2019). Note that for the chosen simulation parameters, the maximum numerical error present in the band gap energy is at most ± 0.1 eV (Kresse, Marsman, and Furthmuller 2018).

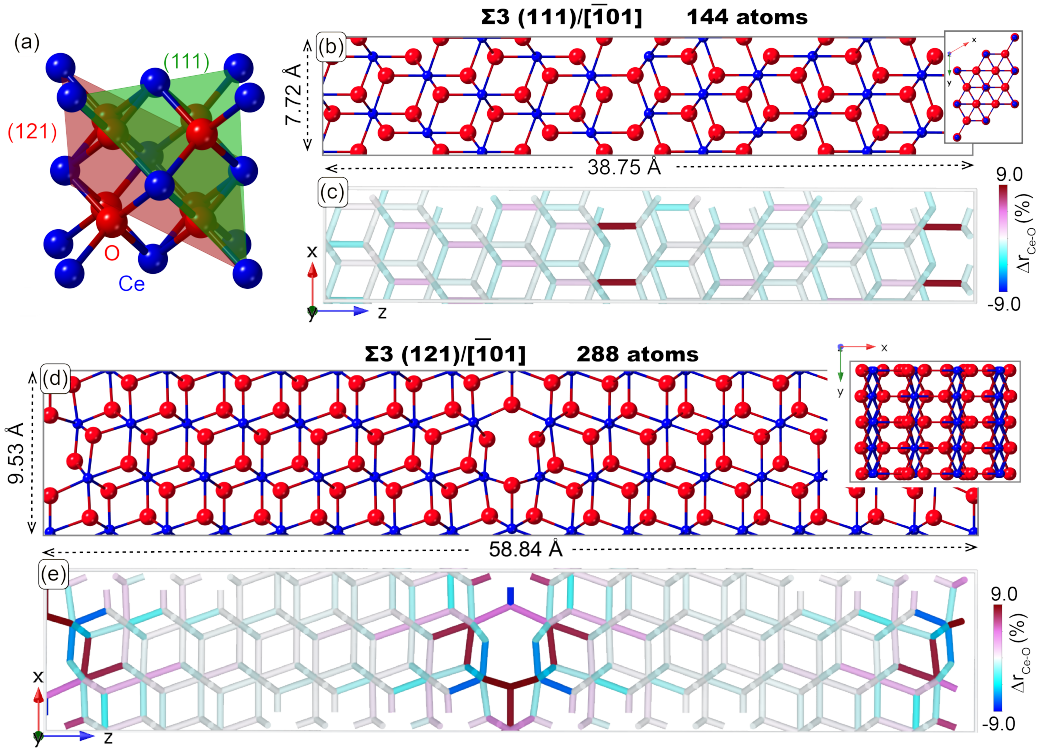

Figure 3.1. (a) Conventional fluorite CeO_2 unit cell. The (121) and (111) lattice planes colored red and green, respectively, indicating the grain boundary planes. Insets show the supercell oriented with grain boundary within the plane–the x-y plane. The xz-plane of the (b) Σ3 (111)/[1̄01] and (d) Σ3 (121)/[1̄01] supercell. The percent deviation of the bonds in the (c) Σ3 (111)/[1̄01] and (e) Σ3 (121)/[1̄01] supercell from the average Ce-O bond length in the supercell. The color map represents $\Delta r_{\mathrm{Ce-O}}$. Blue indicates compressed bonds while red indicates tensile bonds.

A plane wave cutoff energy of 400 eV was used for all cases except for the volume optimization of ceria, where it was set to 520 eV. This cutoff energy was sufficient to converge the forces (Hellman 1937) acting on each ion to 0.01 eV/Å or better. A block Davidson (Broyden 1965) minimization algorithm was used to achieve a convergence in total energy per cell on the order of 0.001 eV or better.

Figure 3.1b and Figure 3.1d are the structure model for the Σ3 (111)/[$\bar{1}$01] and Σ3 (121)/[$\bar{1}$01] grain boundaries, respectively. The grain boundaries were constructed from the conventional fluorite unit cell of CeO_2 using pymatgen (Ong et al. 2013)–an open-source Python library for materials analysis. The choice of grain boundaries was motivated by the large variation in the atomic structure, as well as the presence of high-quality experimental characterization of polycrystalline Ca-doped ceria via electron back-scattered diffraction and transmission electron microscopy (Bowman 2016; William J Bowman et al. 2017). Furthermore, the atomic structure for the Σ3 (111)/[$\bar{1}$01] grain boundary has been experimentally confirmed by B. Feng et al. (2012). The computationally derived CeO_2 Σ3 (111)/[$\bar{1}$01] grain boundary structure agrees well with the high-resolution transmission electron microscopy (HR-TEM) images obtained by B. Feng et al. (2012) and is shown in Appendix Figure A.2. This suggests that the computationally considered structure has a qualitative match with the experimentally measured structure. For the Σ3 (121)/[$\bar{1}$01] grain boundary–a closely related system, YSZ–atomic resolution structural information is also available (Shibata et al. 2003) and is shown in Appendix Figure A.3. The undoped Σ3 (111)/[$\bar{1}$01] grain boundary cell has optimized lattice vectors [7.72, 7.72, 38.75] Å with 144 atoms and was converged with a 3x3x1 gamma-centered k-point grid. The undoped Σ3 (121)/[$\bar{1}$01] grain boundary cell has optimized lattice vectors [9.53, 7.74, 54.84] Å with 288 atoms and was converged

with a 2x3x1 gamma-center k-point grid. A Gaussian smearing with a sigma value of 0.05 eV was employed.

During the initial construction of each grain boundary supercell structure, the inter-grain boundary spacing between respective grains (the z-axis separation) was set to maintain the same cation-anion bond distance across the interfaces as the grain interior. This was motivated by several studies suggesting that ceramic oxides relax to retain a bond length between ions that is similar to the grain interiors (Shibata et al. 2004; Shibata et al. 2002). To minimize the grain boundary interactions between periodic images, the undoped grain boundary cells were constructed from grains with a **c** lattice vector two times the periodic repeat distance of the oriented cell, such that $\mathbf{c} = 2\mathbf{a}_{hkl}$, where hkl are the crystal directions associated with the (111) and (121) interfacial planes. These grain boundary supercells are used for assessing the energy and electronic properties of the undoped and doped grain boundaries. Note that the grain boundary energy difference between the grain boundary supercells, constructed using grains with $\mathbf{c}=\mathbf{a}_{hkl}$ and $\mathbf{c}=2\mathbf{a}_{hkl}$, was 8 meV/Å^2 and 2 meV/Å^2 for the $\Sigma3$ (111)/[$\bar{1}$01] and $\Sigma3$ (121)/[$\bar{1}$01] grain boundaries, respectively.

When doped with an AEM solute, an oxygen vacancy was introduced in the cell to maintain charge neutrality. Pseudopotentials for each AEM solute were chosen such that the total energy was a minimum, and to ensure convergence of the simulations. The O and Ce atoms have been described by $2s^2 2p^4$ and $5s^2 5p^6 6s^2 5d^1 4f^1$ valence electrons, respectively. The valence electrons for Be and Mg were described by $2s^2$ and $3s^2$ while Ca, Sr, and Ba used $3s^2 3p^6 4s^2$, $4s^2 4p^6 5s^2$, $5s^2 5p^6 6s^2$ valence electrons, respectively. All structures considered in this study were subject to full structure optimization.

3.3 Results and Discussion

3.3.1 Grain Boundary Structure and Character

As discussed in Section 1.3 of the Introduction, grain boundaries can be fully described using the 5 macroscopic and 4 microscopic degrees of freedom. Additional atomic-level parameters–like the number of coordination-deficient cation sites, the average cation-anion bond distance, and the grain boundary induced lattice expansion–can further elucidate grain boundary structure-property relationships. A coordination-deficient cation site is a site that has fewer bonds than that of the host cation in the defect-free lattice. Thus for ceria-based compounds, a coordination-deficient cation site will have less than 8 nearest neighbor oxygen atoms. The grain boundary expansion, given by γ_{GB} in Å, is defined as the difference in the z-axis length between the relaxed grain boundary supercell and the corresponding relaxed grain boundary-free supercell divided by two. Hence, γ_{GB} is a measure of the expansion of the pristine ceria lattice vector perpendicular to the grain boundary plane.

Table 3.1 lists the aforementioned atomic-scale parameters and the misorientation angles of the grain boundaries. The $\Sigma 3$ $(111)/[\bar{1}01]$ and $\Sigma 3$ $(121)/[\bar{1}01]$ are both high-angle coincident site lattice boundaries (Bowman 2016)–briefly discussed in Section 1.3 of the Chapter 1–with misorientation angles of 35.26° and 54.74°, respectively. Interestingly, the equidistant (near cubic) polyhedral arrangement of the O ions around the Ce ions tend to remain intact at/near the grain boundary core as can be seen in Figure 3.1. This can be attributed to the large ionicity of the Ce-O bonds. In order to retain the polyhedral arrangement of the host lattice the γ_{GB} is significant, 0.315 Å for $\Sigma 3$ $(111)/[\bar{1}01]$ and 0.471 Å $\Sigma 3$ $(121)/[\bar{1}01]$ grain boundary, in agreement

with experimentally measured values in similar systems (Shibata et al. 2004). Here, it is emphasized that in stoichiometric ceria, the coordination-deficient vacancy sites are structural in origin. The charge neutrality of the compound is maintained for all simulations thus no other point defects were considered to be present at the grain boundary.

Interface-Plane	θ (°)	sites	$\bar{r}^{GB}_{\mathrm{Ce-O}}$ (Å)	γ_{GB} (Å)	ΔE_{GB} (eV/Å^2)
$\Sigma 3$ (111)/[$\bar{1}$01]	35.26	4	2.379	0.315	0.058 (0.93)
$\Sigma 3$ (121)/[$\bar{1}$01]	54.74	4	2.385	0.471	0.093 (1.48)
YSZ $\Sigma 3$ (111)/[$\bar{1}$01] [1.]	35.26	–	–	–	0.031 (0.49)
YSZ $\Sigma 3$ (121)/[$\bar{1}$01] [1.]	54.74	–	–	–	0.037 (0.60)
poly. CeO_2 [2.]	–	–	–	–	0.105 (1.687)

Table 3.1. The interface-plane notation, the misorientation angle θ in °, the total number of coordination deficient cation sites per grain boundary, the average Ce-O bond distance, $\bar{r}^{GB}_{\mathrm{Ce-O}}$ in Å, the z-axis expansion, γ_{GB} in Å, and the grain boundary energy ΔE_{GB} in eV/Å^2 are listed for the two grain boundaries studied in this work. ΔE_{GB} values listed in parenthesis are in J/m^2. Referenced works: 1. Shibata et al. (2004) 2. Zouvelou, Mantzouris, and Nikolopoulos (2008).

The averaged Ce-O bond distance, $\bar{r}^{GB}_{\mathrm{Ce-O}}$, in the $\Sigma 3$ (111)/[$\bar{1}$01] and $\Sigma 3$ (121)/[$\bar{1}$01] grain boundary models are 2.379 Å and 2.385 Å, respectively. These average bond distances in the grain boundaries are practically equal to the bond distances in bulk ceria, $r^{CeO_2}_{\mathrm{Ce-O}} = 2.380$ Å. The excellent agreement between $\bar{r}^{GB}_{\mathrm{Ce-O}}$ and $r^{CeO_2}_{\mathrm{Ce-O}}$, however, does not imply that there are no distortions in the lattice upon introduction of the grain boundary. On the contrary, as shown in Figure 3.1c and Figure 3.1e, up to ±9% bond deviation, $\Delta r^{GB}_{\mathrm{Ce-O}} = \frac{r^{GB}_{\mathrm{Ce-O}} - r^{CeO_2}_{\mathrm{Ce-O}}}{r^{CeO_2}_{\mathrm{Ce-O}}} \times 100$, where $r^{GB}_{\mathrm{Ce-O}}$ is the length of bonds in the grain boundary structure, is observed. Both tensile and compressive strains are present in each grain boundary lattice. The lattice distortions are predominant near

the grain boundary and diminish rapidly a few lattice planes away from the grain boundary.

3.3.2 Thermodynamic Descriptors of Solute Doped Grain Boundaries

In order to compare the stability of ceria in the presence of grain boundaries and solutes, the grain boundary energy ΔE_{GB} is computed using

$$\Delta E_{\mathrm{GB}} = \frac{E_{\mathrm{GB}} - n_{\mathrm{CeO_2}} E_{\mathrm{CeO_2}} - n_{\mathrm{MO}} E_{\mathrm{MO}}}{2\mathrm{A}} \tag{3.1}$$

where E_{GB} is the total energy of the grain boundary supercell with solute M, E_χ is the energy of one formula unit of bulk χ where $\chi = \mathrm{CeO_2}$ or MO listed in Appendix Table A.2, n_χ is the number of formula units of χ in the grain boundary supercell, and A is the area of the xy-plane i.e. the grain boundary containing plane. ΔE_{GB} represents the area normalized excess energy of ceria due to the creation of the grain boundary interface.

As listed in Table 3.1, the ΔE_{GB} of undoped $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary is approximately half the value of the $\Sigma 3$ (121)/[$\bar{1}$01] grain boundary. This is not surprising since the $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary has a high atomic coherency across the interface which is shown in Figure 3.1b and Appendix Figure A.1a. The continuity of the anion and cation sublattices is clearly preserved in the $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary but the $\Sigma 3$ (121)/[$\bar{1}$01] grain boundary has a disruption in the cation sublattice clearly seen in Figure 3.1d and Appendix Figure 3.1b. Other ceramic oxides such as YSZ, display similar dependence of the ΔE_{GB} on the coherency of atoms at the interface (Shibata et al. 2002; Shibata et al. 2003; Shibata et al. 2004).

While experimental and theoretical reports of grain boundary energies in these

systems are scarce, available literature confirms that the grain boundary energies are similar both in value as well as energetic trends (Shibata et al. 2004; Yuan et al. 2016; Zouvelou, Mantzouris, and Nikolopoulos 2008). One study using a multi-phase equilibration technique (Zouvelou, Mantzouris, and Nikolopoulos 2008) determined the grain boundary energy for polycrystalline CeO_2 to be 1.687 J/m^2 at 0 K. This value is reasonably close to the computed values in this work that represent grain boundaries possessing lower energies than one would expect in a polycrystalline ceramic.

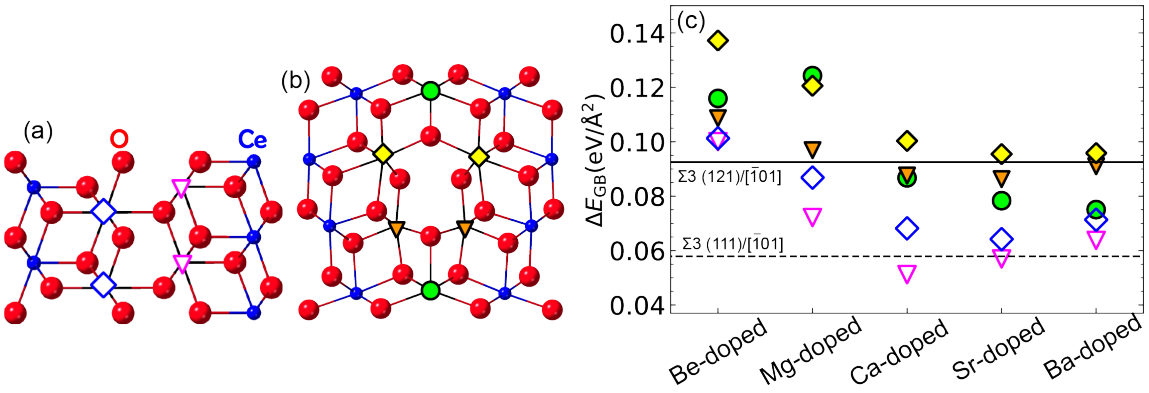

Figure 3.2. The (a) Σ3 (111)/[1̄01] and (b) Σ3 (121)/[1̄01] grain boundary structures with all distinct solute sites indicated by a unique marker-color combination. All ions represented with a marker indicate a core grain boundary cation site which was considered as potential substitutional site. (c) ΔE_{GB} for the Σ3 (111)/[1̄01] (open/dashed markers) and Σ3 (121)/[1̄01] (filled/solid markers) grain boundaries. Each marker corresponds to the ΔE_{GB} for each respective solute site depicted in the grain boundary structure models.

YSZ, a closely related fluorite structured ceramic, has available experimental grain boundary energies (Shibata et al. 2004). In this study the {111} grain boundary was found to possess lower grain boundary energy (∼0.49 J/m^2) while the {121} grain boundary was found to have a higher grain boundary energy (0.6 J/m^2). For the computed grain boundary energies in CeO_2, the values for the undoped grain boundaries follow the same energetic trend with the {111} grain boundary having

the lowest grain boundary energy and the {121} grain boundary having higher grain boundary energy. However, many theoretical calculations do not report grain boundary energy values; for studies which mention grain boundary energies (Yuan et al. 2016), grain boundary energy values around 1 J/m^2 are implied.

The markers in Figure 3.2a and Figure 3.2b illustrate the substitutional sites at the $\Sigma 3$ (111)/[$\bar{1}$01] and $\Sigma 3$ (121)/[$\bar{1}$01] grain boundary of ceria, respectively, where the Be, Mg, Ca, Sr, and Ba solutes are placed to assess their impact on the stability, structure and electronic properties of the lattice. Note that higher concentration of solutes in ceria-based electrolytes at or near grain boundaries (Pratik P Dholabhai et al. 2015; Bokov et al. 2018; Bowman 2016) have been reported both experimentally (Orlovskaya and Browning 2004; W. Lee et al. 2012; William J. Bowman et al. 2015; Lin et al. 2015; William J Bowman et al. 2017; Avila-Paredes and Kim 2006; Y. et al. 2002; Lia et al. 2010; Bowman, Darbal, and Crozier 2020) and theoretically (Mebane and De Souza 2015; Y.-H. Lee et al. 2013; Diercks et al. 2016; Arora and Aidhy 2017) to have higher local concentrations of solutes than usually found in the bulk. Local grain boundary solute concentrations between 20%-40% have been verified by both atom-probe tomography (X. Xu et al. 2020) and atomic resolution transmission electron microscopy (William J Bowman et al. 2017; William J. Bowman et al. 2015; Bowman, Darbal, and Crozier 2020), even when nominal solute concentrations are dilute (0.2%) (X. Xu et al. 2020). By definition, for a cation site to be considered part of the grain boundary core, the site must lie along/on either side of the grain boundary mirror plane and is shown in Appendix Figure A.1. The $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary has two distinct sites, a coordination-deficient site marked by magenta triangles and a fully-coordinated site marked by blue diamonds. The $\Sigma 3$ (121)/[$\bar{1}$01] grain boundary has three distinct sites, the coordination-deficient site marked by

orange triangles and fully-coordinated sites marked by yellow diamonds and green circles. While the sites marked by green circles are fully-coordinated, they favor an asymmetric arrangement of the O-atoms around the site unlike the symmetric cubic arrangement in ceria.

A local grain boundary solute concentration of 25% can be achieved for the $\Sigma 3$ $(111)/[\bar{1}01]$ and the $\Sigma 3$ $(121)/[\bar{1}01]$ grain boundaries by sequentially considering one core grain boundary site, indicated by the markers in Figure 3.2a and Figure 3.2b, within one region of the grain boundary core for doping. A region within the grain boundary core is assumed to have a 2 Å width perpendicular to the grain boundary plane originating from the cation mirror plane and extending towards the bulk. A total of 25 solutes configurations were studied in this work. The large number of atoms in the simulation cell and the rapidly increasing number of configurations prohibit a comprehensive study of other solute concentrations.

Figure 3.2c shows the ΔE_{GB} is greater for the $\Sigma 3$ $(121)/[\bar{1}01]$ grain boundary than the $\Sigma 3$ $(111)/[\bar{1}01]$ grain boundary. Furthermore, for each substitutional site, ΔE_{GB} has a parabolic dependence on the solute cation's ionic radius. The site-dependence of the ΔE_{GB} of the $\Sigma 3$ $(111)/[\bar{1}01]$ grain boundary is low in comparison to that of the $\Sigma 3$ $(121)/[\bar{1}01]$ grain boundary. This can be understood by examining the net bond strain at the solute sites of the grain boundaries shown in Figure 3.1c and Figure 3.1e. In the $\Sigma 3$ $(121)/[\bar{1}01]$ grain boundary, the three distinct solute sites have markedly different net bond strain illustrated by the variation in color in Figure 3.1c, Figure 3.1e, and Appendix Figure A.13-Figure A.14. The net tensile to compressive bond strain ratio is highest in the green-site, intermediate in the yellow site and lowest in the orange site. In comparison, the blue and magenta sites in the $\Sigma 3$ $(111)/[\bar{1}01]$ grain

boundary have similar net bond strains as the Σ3 (121)/[$\bar{1}$01] grain boundary, but much smaller overall net bond strains shown in Figure 3.1c.

For both the grain boundaries, the coordination-deficient cation sites (magenta and orange triangles) are among the lowest energy sites. For the Σ3 (121)/[$\bar{1}$01] grain boundary, the fully-coordinated sites marked by the green circle also have low ΔE_{GB}, especially for the heavier solute cations. These three low ΔE_{GB} sites are also the most strained sites in the grain boundaries. The blue sites in the Σ3 (111)/[$\bar{1}$01] grain boundary and the yellow sites in the Σ3 (121)/[$\bar{1}$01] grain boundary have largest energies displaying a barrier for doping and preference for the Ce-atoms to remain in a site that has coordination and bond-length similar to that of the grain interior. Similar trends in four symmetric tilt grain boundaries have been observed for YSZ (Shibata et al. 2002; Shibata et al. 2004).

It is noteworthy that the addition of Be and Mg make the grain boundaries consistently more unstable across all sites. Apart from the large mismatch in the ionic radii of Ce (R_i = 0.97 Å) with that of Be (R_i = 0.27 Å) and Mg (R_i = 0.57 Å) (Haynes and Lide 2016), the nature of bonding in the native oxides of Mg and Be also dictates the stability of the grain boundary. Unlike the octahedral coordination predominant in Ca (R_i = 1.12 Å), Sr (R_i = 1.42 Å) and Ba (R_i = 1.26 Å) oxides, Be and Mg oxides display a tetrahedral bonding and is denoted in Appendix Table A.2. The Be and Mg solutes relax into interstitial sites to attain this 4-fold coordination where possible, for example, in some of the coordination-deficient sites. The relaxed structure of all solute configurations are presented in Appendix Figure A.13-Figure A.14. Since the Ca solutes have the lowest mismatch in the ionic radii with the host Ce atoms and also more closely match the cubic coordination of the host cation, these solutes render the grain boundary most stable in comparison to the other solutes.

Overall, the ΔE_{GB} critical point appears to be modulated by three primary factors: (a) the local atomic environment of the solute site, (b) the solute size, and (c) the coordination of the solute in its native oxide. The relative difference in ΔE_{GB} between the grain boundaries may be due to the grain boundary packing density. Furthermore, it is evident that the ΔE_{GB} can be more easily modulated by varying the solute type and is more weakly modulated by the substitutional site. Additionally, grain boundary doping strategies attempting to smooth out the potential energy landscape across grain boundaries should focus on Ca or Sr solutes because the lowest grain boundary energies, out of the five solute sites explored, is achieved for the Ca and Sr solutes.

3.3.3 Electronic Structure of Alkaline-Earth Metal Doped Ceria

Aliovalent solutes are often used to increase the number of charge carriers in ceramic oxides (Andersson et al. 2006). Although, they can also introduce localized defect states and/or bands above the band gap activating electronic conduction mechanisms such as polaron hopping (Bishop, Stefanik, and Tuller 2011; Figueiredo and Marques 2013), which can be detrimental to the ionic conductivities. In this section, the AEM solutes are shown to deactivate these potentially detrimental electronic conduction mechanisms by not introducing any defect states above the valence band or in the band gap.

Figure 3.3a and Figure 3.3b show the element-projected density of states (DOS) for the supercells with the $\Sigma 3$ (111)/[$\bar{1}$01] and $\Sigma 3$ (121)/[$\bar{1}$01] grain boundary, respectively. The solid lines show the DOS for the undoped grain boundary and the shaded regions mark the DOS for the solute-doped grain boundary. From the DOS, it is clear that the incorporation of Be and Ba solutes at the grain boundary core does not result in

defect states above the band gap or within it. Similarly, this work finds that none of the solutes impart defect states which can be seen for all solutes in Appendix Figure A.8-Figure A.12.

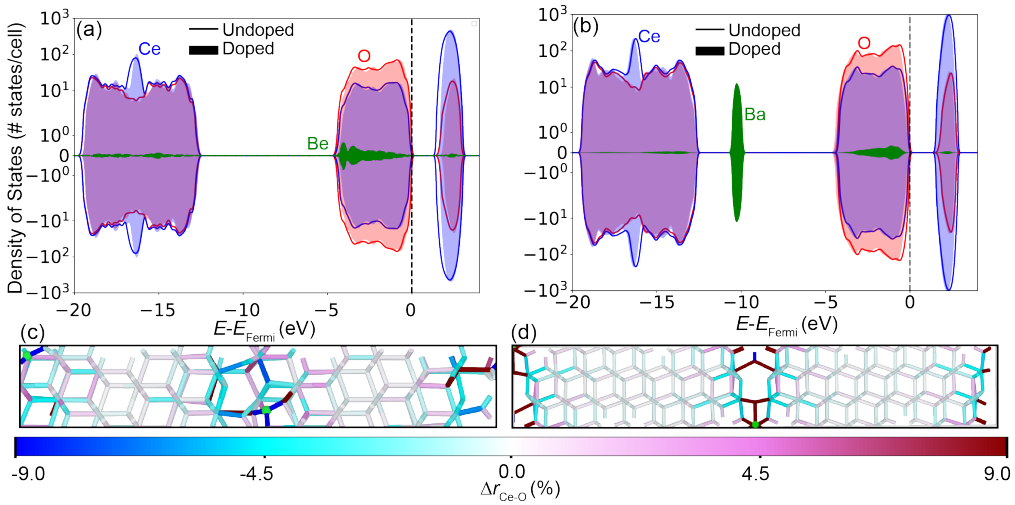

Figure 3.3. The element-projected DOS for the (a) undoped (solid lines) and Be-doped (shaded regions) for the $\Sigma 3$ (111)/[$\bar{1}$01] grain boundaries, and (b) undoped (solid lines) and Ba-doped (shaded regions) for the $\Sigma 3$ (121)/[$\bar{1}$01] grain boundaries. The $\Delta r^{GB}_{\text{Ce-O}}$ of the Be-doped $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary is shown in (c) and that of the Ba-doped $\Sigma 3$ (121)/[$\bar{1}$01] grain boundary is shown in (d). Blue indicates compressed bonds while red indicates tensile bonds. The DOS is shifted such that the top of the valence band is at 0.0 eV.

In the undoped grain boundary, the states at the conduction band maxima (CBM) are dominated by Ce-4f and O-2p states with smaller contributions from Ce-4d and 5p states. The orbital-projected density of states contributions can be seen in Appendix Figure A.9-Figure A.12. The states at the valence band minima (VBM) are mostly O-2p states.

Negligible changes occur in the states present at the VBM and CBM upon doping. For all but the Be-doped grain boundaries, the d, p, and s states of the solute atoms are present at the CBM, resulting in distorted cubic bonding of the solute-O bonds at

the grain boundary; this can be seen Figure 3.3d and Appendix Figure A.9-Figure A.12. For Be-doped grain boundaries, only p and s states of the Be are present at the CBM, indicating a strong propensity of Be to form tetrahedral Be-O bonds as shown in Figure 3.3c.

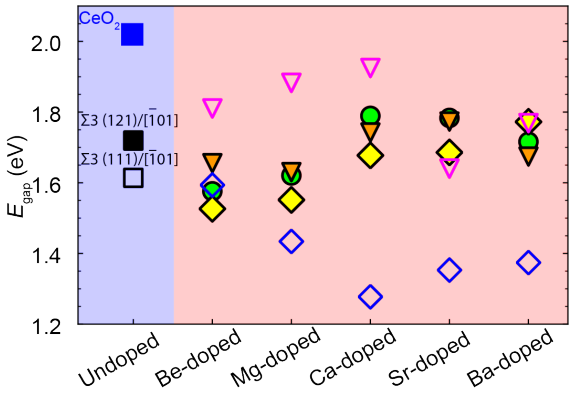

Figure 3.4. (a) The E_{gap} for each solute-grain boundary configuration. The undoped grain boundaries are represented with black squares and the bulk ceria by blue square. The $\Sigma 3$ (121)/[$\bar{1}$01] grain boundary shows minor site and solute dependence while the $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary shows both site and solute dependence.

Relative to bulk ceria, the presence of the grain boundaries and solutes have a significant impact on the band gap as seen in Figure 3.4. The presence of the planar grain boundary defect results in a decrease in the E_{gap}. The calculated E_{gap} values for the $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary and $\Sigma 3$ (121)/[$\bar{1}$01] grain boundary are 1.61 eV and 1.72 eV, respectively. The incorporation of solutes can modulate the band gap further, by up to 0.3 eV relative to the undoped grain boundary. A close inspection of the occupied density of states for AEM doped grain boundaries reveals that states well below the Fermi level alter the Ce-O bonded states in a manner that the E_{gap} decreases with respect to the bulk.

Strain has been shown to strongly alter the band gap of bulk ceria (Ahn et al. 2014; Wen, Lv, and He 2015). Appendix Figure A.6 shows the strain-dependent change

in band gap of bulk ceria. The band gap is found to be linearly dependent on the uniform hydrostatic strain of the bulk ceria lattice. A band gap difference of 0.4 eV is observed for compressive strain of 4% and a -0.4 eV band gap difference for a tensile strain of 4% with respect to the unstrained lattice. A compressive strain results in a band gap increase while a tensile strain results in a band gap reduction compared to the unstrained lattice. Appendix Figure A.6 also shows that band gap difference of 0.05 eV can be resolved in the simulations.

The solute and site dependent band gap variations are found to be about 0.7 eV for the $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary and about 0.2 eV for the $\Sigma 3$ (121)/[$\bar{1}$01] grain boundary. The various solutes in the different sites lead to changes in the lattice parameters of the supercell resulting in volumetric strains of the lattice that are between -1% to 1.5% and are shown in Appendix Figure A.7. However, the lattice strains in the doped grain boundaries are not large enough to result in the observed solute and site dependent band gap variations. For instance, in the case of Ca-doped $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary, a difference of about 0.7 eV is observed for the magenta triangle and blue diamond sites of Ca-doped $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary. As noted in the Appendix Figure A.7a, the volumetric change in the lattice of the Ca-doped $\Sigma 3$ (111)/[$\bar{1}$01] CeO_2 grain boundaries with respect to the unstrained grain boundaries is $\sim$-0.2% for the blue diamond site and $\sim$0.3% for the magenta triangle site. Similar to bulk CeO_2, the magenta triangle site is found to have a band gap increase while the blue diamond site has a band gap decrease due to the compressive and tensile strains, respectively. However, based on the strain-dependent band gap change in the bulk ceria, it is evident that such a small lattice strain cannot be the cause of a $\sim$0.7 eV difference between the band gap for the magenta triangle and the blue diamond sites. In order to make a direct comparison with the undoped $\Sigma 3$ (111)/[$\bar{1}$01] grain

boundary, hydrostatic strain is applied to the undoped Σ3 (111)/[$\bar{1}$01] grain boundary. The band gap changes between these cells demonstrate a similar band gap variations, -0.26 eV for 2% strain and 0.26 eV for -2% strain.

In opposition to the lattice strains, the local strains around the solute site in these two cases are markedly different and an order of magnitude larger. Appendix Figure A.13 shows that the local strain around the magenta triangle site is predominantly compressive, about -5%, and predominantly tensile around the blue diamond site, about 5%-9%. While these calculations cannot prove with certainty that the local strains are the cause of the large site-dependent band gap variation, based on elimination of the lattice-strain and inadequate numerical accuracy as underlying causes, it becomes a possible explanation for the solute-site dependent band gap variation.

For the Σ3 (121)/[$\bar{1}$01] grain boundaries, the band gap changes are well correlated with the average Ce-O bond distances and fluctuate around the undoped Σ3 (121)/[$\bar{1}$01] grain boundary band gap as shown in Appendix Figure A.7. However, for the doped Σ3 (111)/[$\bar{1}$01] grain boundary the band gaps appear to be modulated by three main factors: bond strain, local atomic environment of the grain boundary core, and the ionic radii of the solute atom. The coordination-deficient sites (magenta triangle) in the Σ3 (111)/[$\bar{1}$01] grain boundary have less strain than the fully-coordinated sites (blue diamond) resulting in a linear increase in the band gap shown in Appendix Figure A.13. This trend continues until Sr and Ba where the band gap values decrease, an effect which most likely originates from the increased bond strain that extends well into the bulk. The fully-coordinated sites (blue diamond symbols) are unable to relax since they are sterically hindered by the surrounding anions. The increased strain for these sites increases the hybridization between the Ce $4f$-O $2p$-M nd, np, and ns states where M = Ca, Sr, and Ba and n is the principle quantum number

which decreases the band gap and shown in Appendix Figure A.9. For both the grain boundaries, the Ba-doped ceria maintains a similar band gap compared with the undoped grain boundary samples for all sites considered.

In all, the band gap in ceria is considerably affected by the presence of both planar and solute defects. This can result in heterogeneous electronic properties in experimentally synthesized nanocrystalline ceria. The changes in the band gap due to both solutes and the presence of grain boundaries are correlated with the local atomic structure of the grain boundary, average Ce-O bond distance, and the bond strain. Furthermore, the sensitivity of the electronic structure may be modulated by the grain boundary packing density. For close packed grain boundaries, sites that are sterically hindered may have increased hybridization, thereby decreasing the band gap; whereas unhindered sites show a linear increase in the band gap.

3.3.4 Possible Impact of High Solute Concentrations on Oxide Grain Boundaries

Fundamental studies of grain boundaries in oxide ceramics with the presence of solutes increases the phase-space that needs to be considered, thus, significantly increasing the computational cost and time. The recent advances in computing power makes these much-needed foundational investigations with DFT feasible. This *ab-initio* study on ceria provides critical guidance to larger-length- and time-scale simulations such as molecular dynamics and Monte Carlo simulations (Aidhy, Y. Zhang, and Weber 2013; Aidhy, Y. W. Zhang, and Weber 2014; Arora and Aidhy 2017). Additionally, this lays the foundation for future work exploring relevant grain boundary effects such as space-charge layers (Bondevik, Bjørheima, and Norbya 2020) or migration energy changes (Yuan et al. 2016).

Previously, researchers (B. Feng et al. 2012) have emphasized the importance of investigations that focus on the atomic- and electronic-level properties of different grain boundaries due to their diversity. Currently, the critical factors influencing microscopic properties (such as local bonding, strain and space charge layers) are not understood from an atomic scale perspective. There are many open questions regarding the role and impact that the atomic structure, solute size, local bond strain and composition, as well as the electronic properties play in grain boundary structure-property relationships for this important class of ceramics. Even basic correlations regarding the site-dependent stability of solutes located at grain boundaries was previously unexplored for these systems. Typical grain boundary studies are conducted with interatomic potentials that are fit to bulk properties to gain insights into the grain boundary problem (Aidhy, Y. Zhang, and Weber 2013; Aidhy, Y. W. Zhang, and Weber 2014; Arora and Aidhy 2017).

In ceria, these calculations have shown that grain boundaries possessing less structural coherency have higher grain boundary energies, yet, these energies stabilize when doped with solutes that are larger than the host solute. In general, this is consistent with the principle that coherent interfaces are lower in energy and stable. Solutes that are smaller in ionic radii than the host solute result in a less coherent grain boundary structure and increase the grain boundary energy. Less structural coherency has been previously associated with higher grain boundary energies in fluorite oxides (Shibata et al. 2002; Shibata et al. 2003; Shibata et al. 2004), as well as in similar systems such as YSZ. This trend indicates that ionic materials may exhibit a strong energetic dependence on the structural coherency across the grain boundary. Strategies that aim to decrease the grain boundary energy by doping should focus on solutes larger than the host solute whilst avoiding smaller solutes that render the

grain boundary unstable. A large mismatch between the host and solute ions increases the likeliness of solute segregation and oxygen vacancy binding (Arora and Aidhy 2017) to the grain boundary. Additionally, co-doping with different solutes may be a feasible strategy to optimize both the bulk and grain boundary properties.

The location of the solute within the grain boundary core results in appreciable modulations in the grain boundary energy. However, there is a clear solute size dependence for the grain boundary energy indicating that the type of solute may play a more important role in minimizing the grain boundary energy than the location of the solute. Additionally, the choice of solute, unlike the location of the solute, is accessible via experimental methods. Therefore, it is critically important to carefully select solutes to optimize the grain boundary properties. Future computational work should be directed towards three main areas: (1) identifying solutes that decrease the grain boundary energy, (2) identifying solutes that possess a high grain boundary segregation energy, and (3) possess a low oxygen vacancy binding for ionic conductors.

For applications that depend upon high ionic conductivity, changing the electronic properties via doping is undesirable. Thus, understanding how band gaps are modulated due to the presence of solutes, interfacial strain, and grain boundary character is essential to predict suitable solutes that do not significantly perturb the electronic structure near the band gap. This study demonstrates that the DOS profile around E_{gap} remains unchanged and that the E_{gap} value is perturbed due to the presence of the solutes and grain boundary.

3.4 Summary

In summary, this work utilizes DFT with the GGA+U functional to examine the structure, energetic stability and electronic properties of undoped and AEM doped grain boundaries in ceria. This work studies two high-angle grain boundaries–the $\Sigma 3$ $(111)/[\bar{1}01]$ and $\Sigma 3$ $(121)/[\bar{1}01]$–and find that the $\Sigma 3$ $(111)/[\bar{1}01]$ grain boundary is energetically more stable than the $\Sigma 3$ $(121)/[\bar{1}01]$ grain boundary due to its larger atomic coherency at the grain boundary interface.

To-date, there are no other theoretical studies that exhaustively assess this atomic-scale role of solutes in grain boundaries of ceria, or similar oxide-based systems. Foundational studies such as this are imperative to address questions about the effect of solutes in these well studied, technological relevant materials. This study addresses this need providing a systematic assessment of the local atomic structure, grain boundary geometry, and solutes to determine how these factors modulate grain boundary thermodynamics as well as the electronic properties of the grain boundary.

Considering all the substitutional sites in the grain boundary core, this work demonstrates that when the grain boundaries are doped with $\sim$20% AEM solutes, the grain boundary energies of ceria will depend strongly on the substitutional site's coordination numbers and its local atomic structure. In this work, the lowest energy substitutional sites for each AEM solute is identified showing that Ca, Sr, and Ba solutes stabilize the grain boundaries but Be and Mg solutes render the grain boundaries unstable. The enhancements in the grain boundary stability upon addition of Ca, Sr and Ba can be attributed to similarity in the ionic radii of the solutes and Ce as well as the closely matching coordination of the solute in its native oxide and the ceria lattice. The electronic density of states of doped grain boundaries reveals that no

defect states are present in or above the band gap of the AEM doped ceria, which is highly conducive to maintaining low electronic mobility in these ionic conductors. The electronic properties, unlike the energetic stability, exhibit complex inter-dependence on the structure and chemistry of the host and the solutes. The presence of solutes can modulate the band gap of ceria up to 0.3 eV in comparison to the undoped ceria with grain boundaries.

These results reinforce the claims of Bokov et al. (2018) that solutes can smooth out the energy landscape thereby reducing the grain boundary energy. This smooth potential energy landscape (lower grain boundary energy) could result in an increase in ionic conductivity across grain boundaries and thus a critical factor to consider in optimizing these functional ceramics. Additionally, these results indicate that solutes slightly larger than the host cation stabilize the grain boundary more strongly. Correlations believed to apply in metallic systems such as the grain boundary energy being related to the excess volume at the interface (Wolf and Yip 1993) do not hold for systems that are dominated by ionic bonding, as evident when comparing Figure 3.2c with Appendix Figure A.5 and Appendix Figure A.7 where there is no correlation between the grain boundary expansion or the volume expansion (in any direction) with grain boundary energy.

In the future, advances in computational methods and computing power can enable a comprehensive first-principles based study of more grain boundary structures, solute concentrations as well as the coordinated transport of oxygen-vacancies and ions. This work serves as a guide to these future studies, making an impact on the design of more efficient oxide based ionic conductors.

Chapter 4

ASSESSING COMPOSITIONAL EFFECTS IN CA-DOPED CERIA ON THE CROSS GRAIN BOUNDARY OXYGEN MIGRATION ENERGY

4.1 Motivation

In bulk materials and thin films, defect-grain boundary interactions play a key role in controlling the structural stability, electronic, chemical, and transport properties. Typically, a high-concentrations of aliovalent cations are used in order to introduce charge compensating point defects, thereby increasing the ionic conductivity in these materials. As discussed in Chapter 1, these cations segregate in high concentrations to the grain boundaries. In recent years, there has been an increased interest in the unique properties that nano-crystalline ceramics offer (Cargnello et al. 2013). However, as the surface area of the interfacial regions increase with shrinking grain sizes, the role which high solute concentrations and interfaces play in modulating the overall interfacial materials properties has become of immense importance. These properties can dominate the materials response and deteriorate device performance.

A key challenge in these materials has been the optimization of the cross grain boundary transport of oxygen vacancies where experimental observations have shown a significant reduction in ionic conductivity when compared to the grain. These regions are challenging to optimize due to their minuscule size, significant variations in structure and composition resulting in wide-ranging properties from one grain boundary to another in polycrystalline samples. To overcome the deleterious affects of grain boundaries and improve the overall properties of many polycrystalline electrolytes,

a better understanding of compositional effects on the cross-boundary migration of oxygen vacancies is required.

The cross grain boundary mobility of migrating oxygen vacancies is influenced by two primary factors: 1) thermodynamic and 2) kinetic factors. In polycrystalline ceramics, thermodynamics drives high concentrations of aliovalent solute ions to accumulate at the grain boundary creating a space charge potential region, wherein localized regions can attract, repel, and trap mobile oxygen vacancies. Additionally, the solute ions and oxygen vacancies have a tendency to form defect associations that require energy to separate oxygen vacancies from the solute ions. Kinetically, the migration of oxygen vacancies near the solute ions can be reduced, however, whether this effect occurs within the disordered grain boundary regions remains unclear.

In this chapter, using DFT calculations, the impact that high local solute concentrations in the grain boundary core is systematically investigated assessing the relative segregation energy and migration energy of oxygen vacancies in the near grain boundary regions of the $\Sigma 3$ (111)/[$\bar{1}$01] in Ca-doped ceria. These results provide critical insight into the role that high local solute concentrations play in modulating the migration energy of oxygen vacancies improving the understanding of cross boundary oxygen transport directing further grain boundary studies aimed to engineer grain boundaries with improved ionic conductivity.

4.2 Computational Methods

The same parameters used in Chapter 3 are used throughout these calculations. A plane wave cutoff energy of 400 eV was sufficient to converge the forces (Hellman 1937) acting on each ion to 0.01 eV/Å or better. A block Davidson (Broyden 1965)

minimization algorithm was used to achieve a convergence in total energy per cell on the order of 0.001 eV or better. A gamma center k-point mesh grid density of 3x3x1 was used for all calculations. The 2x2x2 bulk supercell of ceria was created from the conventional unit cell of ceria to calculate the bulk migration energy of an oxygen vacancy. For these structures, a gamma center k-point mesh grid density of 3x3x3 was used for all calculations.

The nudged elastic band method (NEB) (Henkelman, Arnaldsson, and Jønsson 2006) was used to calculate the migration energy of oxygen. The spring constant was set to -5 with three NEB images per migration event. The forces were set to less than 0.05 eV/Å under constant volume to reduce the computational cost of the large supercell sizes. The localized state of the electrons is confirmed using partial charge density. To confirm migration energy barrier was converged with three images, a second calculation with five images was performed with the bulk supercell of ceria.

4.3 Oxygen Vacancy Segregation in Undoped and Ca-Doped Σ3 (111)/[$\bar{1}$01] Grain Boundaries

The interplay between the aliovalent solutes and charge compensating oxygen vacancies can results in increased or decreased ionic conductivity originating from the sensitive nature of defect-defect interactions to variations in solute concentrations (Pratik P. Dholabhai et al. 2010a; Dholabhai and Adams 2012; Pratik P Dholabhai et al. 2015; Aidhy, Y. Zhang, and Weber 2013; Aidhy, Y. W. Zhang, and Weber 2014; Aidhy 2016). Grain boundary populations have significant compositional variations owing to the non-uniformity in grain boundary character and thermodynamic stability which can cause heterogeneous segregation of cations to grain

boundaries. Both in the bulk and at the grain boundary, the mobile oxygen vacancies can be sequestered by the solute ions thereby decreasing the ionic conductivity in these regions. Optimization of the defect-defect interactions between solutes and oxygen vacancies has been widely studied in the bulk while the study of such compositional effects on the cross grain boundary migration of oxygen vacancies has not. In this section, the energetic stability of oxygen vacancies and resulting electronic structure of the near grain boundary region is explored with and without the presence of high local Ca solute concentrations.

4.3.1 Atomic Structure of Undoped and Ca-Doped $\Sigma 3$ (111)/[$\bar{1}$01] Grain Boundary

To explore composition dependent variations in the oxygen vacancy stability in the near grain boundary regions, four oxygen planes are investigated as potential host planes for introducing an oxygen vacancy in the undoped and Ca-doped $\Sigma 3$ (111)/[$\bar{1}$01] ceria grain boundary. The undoped $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary structure model is shown in Figure 4.1a and the Ca-doped $\Sigma 3$ (111)/[$\bar{1}$01] grain boundaries are shown in Figure 4.4a and Figure 4.5a which represents the two unique solute locations for this grain boundary structure discussed in Chapter 3. The vertical black dashed lines delineate each oxygen plane i.e. the plane where the oxygen vacancy site will be incorporated. Figure 4.1b illustrates the band projected charge density for the undoped $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary with an oxygen vacancy. The orange isosurfaces are the real-space representation of the localized Ce $4f$ electrons. The numbers represent the average Ce-O bond distance for that Ce ion where the Ce which contain the localized electrons have a larger Ce-O bond distance.

The crystal direction with the lowest oxygen vacancy migration energy in ceria is

along the <001> crystallographic direction in the bulk and the grain boundary (Yuan et al. 2016) regions and this direction is indicated in Figure 4.1a. As the ultimate goal is to understand variations in the oxygen vacancy migration energy in the near grain boundary region, an oxygen vacancy is only moved to nearest neighbor (NN) oxygen sites along this family of crystal directions for each oxygen vacancy location. Oxygen plane 1 was found to be the most energetically stable oxygen vacancy location for the undoped and Ca-doped grain boundaries, consistent with previous studies (Yuan et al. 2016; Aidhy, Y. W. Zhang, and Weber 2014).

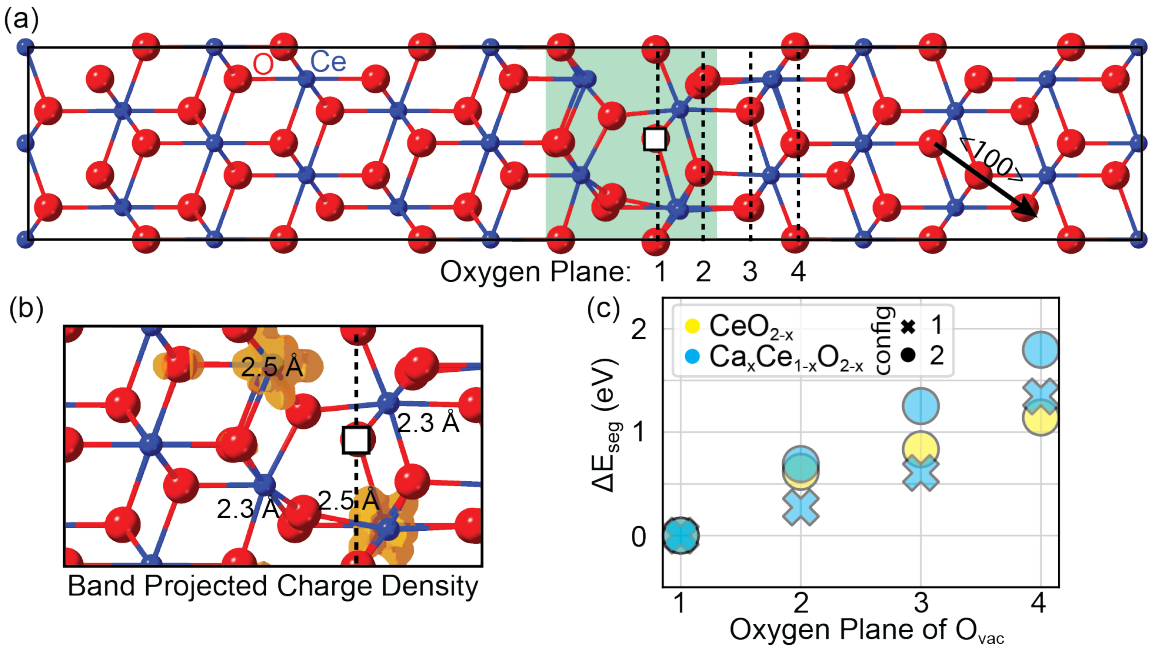

Figure 4.1. (a) Undoped Σ3 (111)/[$\bar{1}$01] grain boundary structure model. Dashed vertical lines illustrate each oxygen plane. The oxygen vacancy is indicated by the square. The transparent green rectangle illustrates the grain boundary core region. (b) Band decomposed charge density isosurface (orange) for the two localized Ce 4f electrons near the oxygen vacancy. (c) Segregation energy for the undoped (●) and Ca-doped (●) Σ3 (111)/[$\bar{1}$01] when the oxygen vacancy is located at each plane shown in (a). Red circles indicate oxygen ions and blue circles indicate cerium ions.

4.3.2 Vacancy Segregation Energy within each Oxygen Plane

Oxygen vacancies are thought to be attracted to grain boundary regions due to the discontinuity introduced by the interface formation resulting in increased energy and high concentrations of solute ions. However, it is unclear what role these additional solute ions play in modulating the energy landscape experienced by the oxygen vacancies with respect to their segregation and cross grain boundary migration. To determine the influence of Ca solutes in the near grain boundary region and assess their impact on the segregation of oxygen vacancies, the oxygen vacancy segregation energy at four oxygen planes–labeled in Figure 4.1a–is determined for the undoped and Ca-doped $\Sigma 3$ (111)/[$\bar{1}$01] grain boundaries.

Yuan et al. (2016) found the segregation behavior of oxygen vacancies to the $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary to be influenced primarily by the first three oxygen planes i.e. the oxygen vacancy migration energy is modulated more strongly within the first 3-4 oxygen planes of the grain boundary core. Motivated by these findings, this study investigates only the near grain boundary planes to assess the impact of Ca solute ions. Figure 4.1c shows the segregation energy for the undoped $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary with the yellow circle markers and the Ca-doped $\Sigma 3$ (111)/[$\bar{1}$01] grain boundaries with the cyan colored markers. For the undoped $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary, the relative change in the segregation energy within the first four oxygen planes agree very well with those calculated by Yuan et al. (2016) for the same grain boundary.

The relative segregation energy, given by $\Delta E_{\text{seg}} = E^{i}_{\text{vac}} - E^{\text{min}}_{\text{vac}}$, represents the energy that must be provided to the system to move the oxygen vacancy from the most stable location. Therefore, the lower the segregation energy the more likely an

oxygen vacancy is to inhabit that site. The larger the segregation energy the more energy needed to move an oxygen vacancy from the most stable oxygen plane. The quantity $E_{\text{vac}}^{\text{min}}$ represents the system energy where the oxygen vacancy is located in the most stable plane and E_{vac}^{i} is the system energy when the vacancy is located in plane i.

When comparing the relative change in the segregation energy of the first three oxygen planes in the undoped grain boundary to the first ($\mathbf{x}$) Ca-doped grain boundary, this system has smaller variations between each oxygen plane. This would imply the oxygen vacancy is able to move more freely within the grain boundary core or the oxygen vacancy location is less selective within these layers. However, as the oxygen vacancy moves into the 3NN position from the Ca ion, the segregation energy increases above the undoped grain boundary. This increase in segregation energy is likely due to the separation of the defect associate.

The most notable impact the presence of the Ca solute ion has on the relative segregation energy in the first Ca-doped system compared to the undoped system occurs within the second and third oxygen planes. These planes represent the 1NN and 2NN shells to the Ca ion. Note, the first two oxygen planes are within the 1NN shell, the third oxygen plane is in the 2NN shell, and the fourth oxygen plane is in the 3NN shell.

The 3NN shell (oxygen plane 4) possess the largest increase in segregation energy off all sites. The increased energy is similarly observed in bulk systems where the separation of the vacancy from the solute ion increases the total system energy (Minervini, Zacate, and Grimes 1999). Additionally, around the 3NN shell is where the effects of the solute ion–such as changes in the migration energy (Pratik P. Dholabhai

et al. 2010a)–are thought to tapper off and approach bulk like values (Y.-L. Lee et al. 2017).

Looking at the second (●) Ca-doped system, the relative segregation energy has larger variations between each oxygen plane compared to both the undoped and first Ca-doped system. However, the increase in energy from one plane to the next is remarkably consistent around ∼0.6 eV. Interestingly, unlike the first Ca-doped system there is no sharp increase in segregation energy. More surprising is the variation between the two Ca-doped systems segregation energy where the more energetically stable second Ca-doped system displays less variability and higher segregation energies than the slightly higher energy first Ca-doped system. This seems to indicate that the site the solute ion sites at in the grain boundary core would impact the segregation of oxygen vacancies more. The second Ca-doped system has far fewer unique characteristics and variations at the different oxygen planes compared to the other 2 systems.

Overall, the impact of high solute concentrations at the grain boundary does not significantly alter the relative segregation energetics of the mobile oxygen vacancies in the near grain boundary region. However, whether the segregation energy of the oxygen vacancy is decreased in comparison to the undoped system appears to be dependant on the location of the solute within the grain boundary core.

4.3.3 Electronic Properties along the Migration Path

The orange isosurfaces in Figure 4.1b illustrates the partial (band decomposed) charge density summed over all k-points for the two defect states created by incorporating an oxygen vacancy in the undoped ceria lattice. Unlike the partial charge density for bulk ceria, the two defect states at the grain boundary are not fully localized on

the Ce ions. Interestingly, the orange isosurface of the Ce ion to the left of the dashed line in Figure 4.1b shows the charge density is partially delocalized on the nearby O ion while the charge density is fully localized for the Ce ion to the right of the dashed line. As the oxygen vacancy migrates towards the bulk, one of the electrons localized on the Ce $4f$ ion follows the oxygen vacancy while the the second electron remains localized at the grain boundary core, shown in Appendix B Figure B.1-B.2. Appendix B discusses challenges and observations regarding these calculations which may be useful towards future grain boundary studies in this/similar systems.

Future studies should focus on the inclusion of additional oxygen vacancies to ascertain how oxygen vacancy-defect interactions influence the stability and segregation behavior of oxygen vacancies in the nearby grain boundary region. Additionally, other aliovalent solute ions such as Sr and Ba–which can stabilize the grain boundary energy (Boland et al. 2021)–may exhibit similar segregation phenomenon for oxygen vacancies and provide solute size trends on oxygen vacancy segregation. Lastly, the importance of investigating other grain boundaries can not be overlooked as much experimental characterization has shown grain boundary properties can vary significantly from one grain boundary to another.

4.4 Bulk and Cross Grain Boundary Oxygen Vacancy Migration Energy in Ca-Doped Σ3 (111)/[$\bar{1}$01]

The migration of oxygen vacancies through the bulk with and without the presence of solute ions has been well studied (De Souza, Ramadan, and Hörner 2012; Koettgen et al. 2018; Sun, Fu, and Yang 2018; Nolan and Watson 2005; H. Shi et al. 2020; Schmitt et al. 2020). Recent efforts have shifted the focus from optimization of

the grains to the grain boundary regions to further improve the properties of these electroceramics. In this section, the results for the bulk oxygen vacancy migration energy of ceria is compared with values obtained in literature then the migration energy path for an oxygen vacancy is explored in the near grain boundary region for the Ca-doped $\Sigma 3$ (111)/[$\bar{1}$01].

4.4.1 Oxygen Vacancy Migration Energy in Bulk Ceria

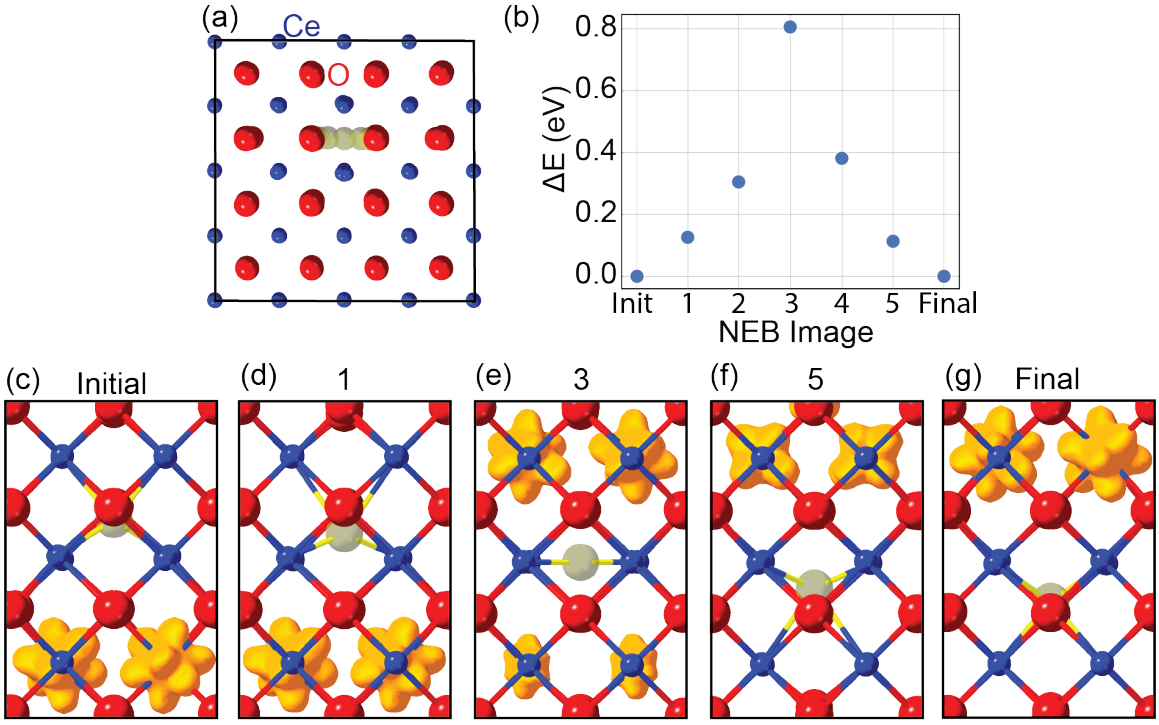

Figure 4.2. (a) 2x2x2 supercell of ceria with one oxygen vacancy. Translucent yellow circles indicate intermediate state positions for the oxygen ion during the NEB calculation. Red circles indicate oxygen ions and blue circles indicate cerium ions. (b) Oxygen vacancy migration energy barrier calculated using NEB. (c-g) The orange isosurface illustrates the band projected charge density for the localized Ce $4f$ electrons during each image during NEB calculation.

Figure 4.2a shows a bulk 2x2x2 supercell of ceria with one oxygen vacancy where

the translucent ions represent the intermediate state positions of the migrating oxygen ion to compute the migration energy in bulk ceria. The migration energy was computed using NEB with a transition state energy of 0.81 eV. Both 3 and 5 images were used with no difference in the barrier height observed where Figure 4.2b shows the migration energy barrier.

It is interesting to note as the oxygen vacancy migrates from the initial to final state position shown in Figure 4.2c-g, the charge at the mid-point (image 3) is partially localized between the 2 sets of 2 Ce ions in which the electrons localize onto in the initial and final state images. Despite the symmetric nature of the migration event, the charge density does not reflect this. The NEB images from the initial state to image 2 appear well localized while states 4 and 5 are not as well defined. This asymmetry could be one reason the NEB energy is not perfectly symmetric as one would expect. One method to fix this issue would be to specify the magnetic moment (using the MAGMOM tag) of each Ce ion in each NEB image forcing a symmetric arrangement of the electrons for the NEB calculation. This asymmetric NEB path was not observed for the NEB calculation with 3 images and is shown in Appendix Figure B.3. Additional discussion regarding these calculations can be found in the NEB section in Appendix B.

Figure 4.3a shows the band structure for the initial state of the NEB calculation demonstrating the two defect bands that localized on the nearby Ce ions at/below the Fermi level. Figure 4.3b-c shows 2 representative element projected DOS for the NEB migration path of bulk defective ceria with Ce in blue, O in red, and the total DOS in green. Both the DOS and band structure show the two defect states. As the oxygen vacancy migrates from one site to the next the two defect states transition from below the Fermi level to above.

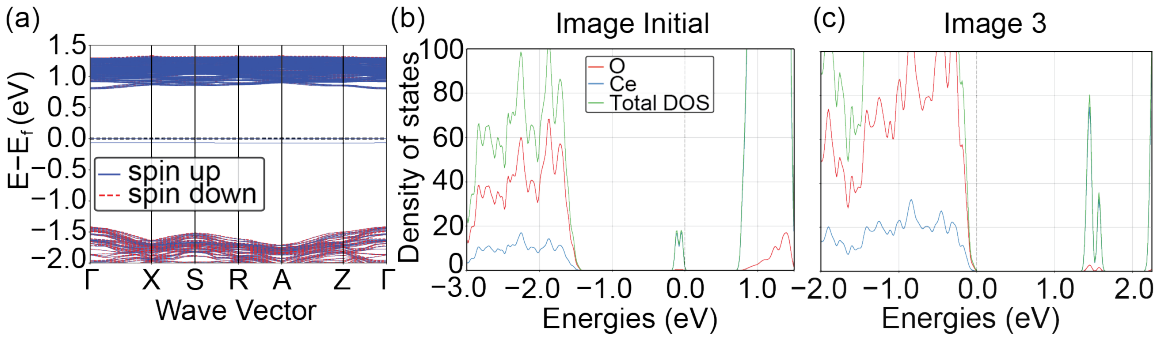

Figure 4.3. (a) Band structure of the initial state bulk defective ceria with spin up (blue) and down (red) plotted. The localized Ce $4f$ bands are near the Fermi level. (b) Initial state element projected DOS for bulk defective ceria showing. (c) Element projected DOS for bulk defective ceria during oxygen vacancy migration.

4.4.2 Oxygen Vacancy Migration Path in Ca-Doped Σ3 (111)/[$\bar{1}$01] Ceria Grain Boundary

Three oxygen vacancy hops forming a migration path away from the grain boundary core (labeled I, II, and III in Figure 4.4a and Figure 4.5a) are investigated in the Ca-doped Σ3 (111)/[$\bar{1}$01] grain boundaries. The oxygen vacancy migrates from oxygen plane: 1 to 2, 2 to 3, and 3 to 4 in each structure and the oxygen ion in the reverse direction. Each hop consists of an initial state image labeled A, intermediate state images labeled B-D, and a final state image labeled E. The relaxed structure model of oxygen vacancy hop I in the first Ca-doped grain boundary system labeled according this the scheme is shown in Figure 4.4b. Figure 4.4c and Figure 4.5b represent the energy of the initial-intermediate-final state(s) during the migration of the oxygen vacancy.

Figure 4.4c shows the oxygen vacancy migration energy barrier near the Ca solute ion as the oxygen vacancy migrates along the path. The horizontal black arrows indication the barrier height as the oxygen vacancy moves away from/towards the

grain boundary. The migration energy towards the grain boundary core is significantly reduced compared to the bulk migration energy of 0.81 eV. The migration energy towards the grain boundary is E_{21}^{I}=0.15 eV, E_{32}^{II}=0.30 eV, and E_{43}^{III}=0.31 eV.

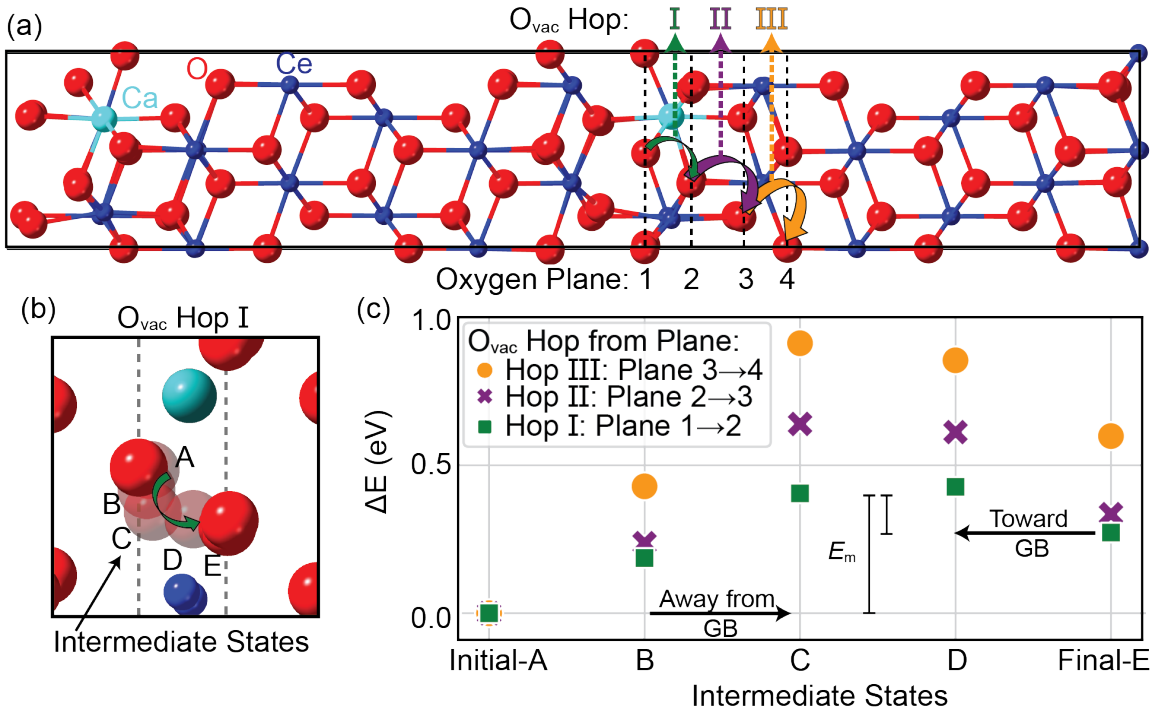

Figure 4.4. (a) Ca-doped Σ3 (111)/[$\bar{1}$01] grain boundary structure model. The black dashed lines indicate oxygen planes 1-4 where the oxygen vacancy is located in plane 1. Colored arrows indicate a migration event where the oxygen ion and vacancy exchange locations along <001>. (b) The NEB intermediate states associated with oxygen vacancy hop I. The oxygen vacancy's initial state and final state are located in plane 1 and 2, respectively. Translucent red circles indicate intermediate state positions of the oxygen ion for the initial state to the final state (A-E) and the green arrow indication the oxygen vacancy migration direction. (c) Energy associated with each initial-intermediate-final state(s) calculated using the NEB method for oxygen vacancy hop: I (■), II (✖) and III (●). Cyan circles indicate Ca ions otherwise the color scheme is the same.

Figure 4.5b shows the oxygen vacancy migration energy barrier for the second system. This system also has a reduction in the barrier height compared to the bulk migration energy and is significantly lower compared to the first Ca-doped system.

The migration energy towards the grain boundary is E_{21}^{I}=0.04 eV, E_{32}^{II}=0.17 eV, and E_{43}^{III}=0.18 eV. This decrease in the barrier height as the oxygen vacancy migrates towards the grain boundary is not to surprising due to the lower energy of the system as the oxygen vacancy migrates towards the grain boundary core and the Ca solute.

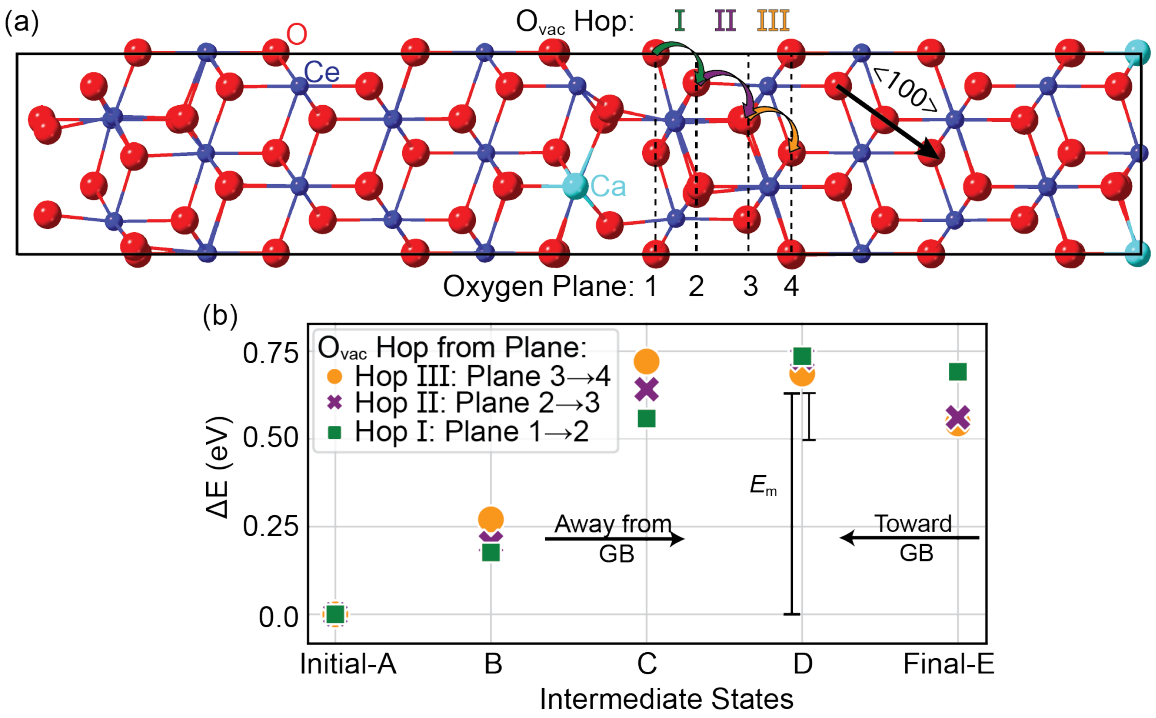

Figure 4.5. (a) 2^{nd} Ca-doped Σ3 (111)/[$\bar{1}$01] grain boundary structure model with vertical dashed lines indicating each oxygen plane. The oxygen vacancy is located in plane 1. (b) Energy associated with each initial-intermediate-final state(s) calculated using the NEB method for oxygen vacancy hop: I (■), II (✖) and III (●).

Both from a thermodynamics and kinetics standpoint, the migration of oxygen vacancies is not inhibited towards the grain boundary as the migration energy values are significantly lower in all cases than the bulk migration energy. However, the models lack adequate complexity to account for the interaction of multiple oxygen vacancies that would inevitably be attracted to the grain boundary core due to the increase stability of the oxygen vacancies and the presence of the solute ions. Since oxygen

vacancies are known to repel one other it is likely that the presence of additional vacancies at/near the grain boundary core would be restricted preventing additional oxygen vacancies from migrating to oxygen sites within the 1NN shell of another oxygen vacancy.

The migration energy away from the grain boundary core for the first Ca-doped system is E_{12}^{I}=0.43 eV, E_{23}^{II}=0.64 eV, and E_{34}^{III}=0.91 eV. The first two migration energy barriers are significantly lower than the bulk migration energy. The decreased oxygen vacancy migration energy within these layers could result in more mobile oxygen vacancies within the grain boundary core. The increased migration energy barrier between the oxygen plane 3 and 4 would certainly create a kinetic barrier for oxygen vacancy migration. Additionally, this increased barrier height is accompanied by the largest increase in the relative segregation energy. These two factors seem to indicate that the migration of oxygen vacancies away from the grain boundary core is both thermodynamic and kinetic unfavorable. It is therefore not hard to imagine that these factors would create an overall decrease the grain boundary conductivity which may explain the loss in conductivity seen at the grain boundary.

The increase in migration energy as the oxygen vacancy migrates away from the Ca ion can be understood by considering the affinity for doped oxides to form defect associations and the lower energy for oxygen vacancies at the grain boundary. Defect associates are electrically neutral clusters ($\mathrm{Ca}_{\mathrm{Ce}}^{\text{“}}$-$\mathrm{V}_{\mathrm{O}}^{\cdot\cdot}$)–composed of a negatively charged cations ($\mathrm{Ca}_{\mathrm{Ce}}^{\text{“}}$) and positively charged vacancy sites ($\mathrm{V}_{\mathrm{O}}^{\cdot\cdot}$) with respect to the pristine host lattice–which are energetically more stable (Koettgen et al. 2018), therefore, attractive in nature resulting in defect clusters. The act of separating these defect associates requires additional energy causing an increase in the ground state energy as well as the migration energy as the defect association is separated.

Surprisingly, in the second Ca-doped system the migration energy barrier away from the grain boundary core for each hop has an almost identical value at 0.73 eV. The only significant difference between each barrier height is the image at which the barrier height peaks.

Yuan et al. (2016) calculated the migration energy barrier for the undoped $\Sigma 3$ $(111)/[\bar{1}01]$ grain boundary finding in the near grain boundary region the migration energy was lower going towards the grain boundary and larger away from the grain boundary. Comparing the migration energies for the 2 Ca-doped system with those computed in the work by Yuan et al. (2016) for the undoped $\Sigma 3$ $(111)/[\bar{1}01]$ grain boundary, the Ca-doped migration energies away from the grain boundary for the first system are 0.30 eV-0.50 eV smaller than the undoped case while the second system has similar values to the undoped system. However, the migration energy values towards the grain boundary in general are significantly smaller except for hop III in the first Ca-doped system. It should be noted the work was calculated using a different functional and +U value so the absolute values would likely not be identical.

The migration energy values calculated with the computational parameters in this work for hop I from the undoped grain boundary are similar to those reported in Yuan et al. (2016). The migration energy towards the grain boundary, however, is lower. Currently only the hop I has been calculated for the undoped grain boundary due to issues converging the NEB calculations. Appendix B discusses additional challenges encountered during this project and solutions to manage such issues when they do occur.

4.5 Summary

In summary, using DFT calculations, the impact of high local solute concentrations in the grain boundary core is systematically investigated assessing the segregation energy and migration energy of oxygen vacancies in the near grain boundary regions of the $\Sigma 3$ (111)/[$\bar{1}$01] in Ca-doped ceria. The presence of Ca solute ions at the grain boundary does not significantly alter the relative segregation energetics of the mobile oxygen vacancies in the near grain boundary region compared to the undoped grain boundary. However, the segregation behavior of oxygen vacancies is sensitive to the site of the solute ion in the grain boundary. For these two Ca-doped systems both an increase and decrease in oxygen vacancy segregation was observed.

The most significant impact of the Ca solute ions at the grain boundary is the reduction in the oxygen vacancy migration energy in the grain boundary core. Depending on the oxygen plane and solute ion location, the presence of Ca solute ions at the grain boundary can result in similar, smaller, and larger migration energy barriers both away from and towards the grain boundary compared with the undoped grain boundary (found in literature) and bulk ceria. Overall, for a majority of the oxygen planes, the presence of the Ca solute ion does decrease the oxygen vacancy migration energy. However, despite the reduced migration energy barriers, the presences of kinetically limited oxygen planes would likely dominate the materials response decreasing the cross grain boundary conductivity. Additionally, the apparent sensitivity of the migration energy to the solute site could result in some regions along the grain boundary plane being more or less conducive for oxygen vacancy transport. Future studies could focus on the effects of Ca solutes both in the bulk and grain boundary core (more realistic doping schemes) as well as the effects of additional

oxygen vacancies which would significantly impact the allowed migration pathways across the grain boundary.

While the segregation of oxygen vacancies to the grain boundary is still favorable both thermodynamically and kinetically, the presence of Ca solute ions dramatically alters the kinetics in the near grain boundary region. However, the limitations of DFT preclude investigations into the effects of multiple oxygen vacancies and realistic solute decorating schemes within the bulk and grain boundary core, due to the large computational cost associated with DFT and the configuration space. Despite this, these results provide critical insight into the role that high local solute concentrations play in modulating the migration energy of oxygen vacancies improving the understanding of cross boundary oxygen transport.

Chapter 5

COMPUTATIONAL SYNTHESIS OF 2D MATERIALS: A HIGH-THROUGHPUT APPROACH TO MATERIALS DESIGN

5.1 Motivation

As discussed in Chapter 1 Section 1.4, the emergence of atomically thin, single-layer graphene spawned a new class of two-dimensional (2D) materials (M. Xu et al. 2013; Novoselov 2011) with an extraordinary range of properties - from large band-gap insulators to the very best conductors, the mechanically tough to soft and malleable, and semi-metals to topologically insulating (Singh, Mathew, et al. 2015; Paul et al. 2017; Blonsky et al. 2015; Akiyama, Kawamura, and Ito 2021). However, large-scale, low-defect, and reproducible synthesis of 2D materials remains a challenging task. Substrate-assisted methods, such as chemical vapor deposition, are a promising avenue to realize large-area, low-defect flakes (Novoselov et al. 2012) of various 2D materials–both with and without a van der Waals (vdW) bonded bulk counterpart.

To enable the functionalization of 2D materials and to assist in the selection of substrates for synthesis, a detailed understanding of the substrate-assisted modification of energetic, physical, and electronic properties of 2D materials is required. This chapter discusses the development and utilization of an open-source high-throughput workflow package, *Hetero2d*, that searches for low-lattice mismatched substrate surfaces for any 2D material. The *Hetero2d* workflow package is inspired by existing community workflow packages (Jain et al. 2013; Jain et al. 2015; Mathew et al. 2016; Mathew

et al. 2017; Ong et al. 2013) but tailored to address scientific questions regarding the stability and properties of 2D-substrate heterostructured materials.

Hetero2d provides automated routines for the generation of low-lattice mismatched heterostructures for arbitrary 2D materials and substrate surfaces, the creation of vdW-corrected density-functional theory (DFT) input files, the submission and monitoring of simulations on computing resources, and the post-processing of the key parameters. Currently, the post-processing computes: (a) the interface interaction energy of 2D-substrate heterostructures (energetics), (b) the identification of substrate-induced changes in the interfacial structure (structural), and (c) charge doping of the 2D material (DOS and charge density). The 2D-substrate information generated by *Hetero2d* routines is stored in a MongoDB database tailored for 2D-substrate heterostructures.

As an example, the capability of *Hetero2d* is demonstrate by identifying stable 2D-substrate heterostructures for four 2D materials, namely 1H-MoS_2, 1T- and 1H-NbO_2, and hexagonal-ZnTe, considering 50 cubic elemental substrates on low-index surface planes. The substrates that sufficiently stabilize the formation energies of these 2D materials are found to be Cu, Hf, Mn, Nd, Ni, Pd, Re, Rh, Sc, Ta, Ti, V, W, Y, and Zr and possess binding energies in the range of $\sim$0.1-0.6 eV/atom. Upon examining the thickness changes in the post-adsorbed 2D material, the charge transfer, and the electronic density of states at the 2D-substrate interface, a covalent type of bonding is found at the interface which suggests that these substrates can be used as contact materials for the 2D materials. Hetero2d is available on GitHub as an open-source package under the GNU license.

5.2 DFT Approach to Identifying Stable 2D-Substrate Heterostructures

2D materials are inherently meta-stable materials and are often created by peeling 2D films from layered, vdW bonded bulk counterparts. Their meta-stability arises from the removal of the vdW bonds between the individual flakes. However, the vdW bonds are an order of magnitude weaker than the in-plane covalent or ionic bonds of 2D materials, thus many 2D materials can remain stable at room temperature or above. A quantitative measure of the stability of 2D materials to remain as a free-standing 2D film is given by the formation energy, ΔE^f_{vac}, with respect to the bulk phase

$$\Delta E^f_{\text{vac}} = \frac{E_{\text{2D}}}{N_{\text{2D}}} - \frac{E_{\text{3D}}}{N_{\text{3D}}}, \tag{5.1}$$

where E_{2D} is the energy of a 2D material in vacuum, E_{3D} is the energy of the bulk counterpart of the 2D material, and N_{2D} and N_{3D} are the number of atoms in the unit cell of 2D and bulk counterpart, respectively. The only requirements that the bulk counterpart of the 2D material (the 3D phase) must meet are that the bulk material must have the same composition as the 2D material and it must be the lowest energy phase. If a 2D material does not have a corresponding 3D phase, the energy of the 3D phase can be taken from the sum of bulk phases normalized to equal the elemental composition of the 2D material.

The ΔE^f_{vac} of a 2D material indicates the stability of a 2D flake to retain the 2D form over its bulk counterpart, where the higher the ΔE^f_{vac}, the larger the driving force to lower the free energy. Singh, Mathew, et al. (2015) and others have shown that when the $\Delta E^f_{\text{vac}} < 0.2$ eV/atom, the 2D materials are stable as a free-standing film, but for larger ΔE^f_{vac}'s they are highly unstable and may only be synthesized using substrate-assisted methods (Singh, Mathew, et al. 2015; Haastrup et al. 2018).

For substrate surfaces to stabilize a 2D material during the growth processes, the 2D-substrate heterostructure should be energetically stable. Thus the interactions between the 2D material and substrate surface have to be attractive in nature. This interaction energy known as the binding energy can be estimated as, $\Delta E_{\mathrm{b}} = (E_{\mathrm{2D}} + E_{\mathrm{S}} - E_{\mathrm{2D+S}})/N_{\mathrm{2D}}$, where $E_{\mathrm{2D+S}}$ is the energy of the 2D material adsorbed on the surface of a substrate, E_{S} is the energy of the substrate slab, E_{2D} is the energy of the free-standing 2D material, and N_{2D} is the number of atoms in the unit cell of the 2D material. Note, strain is applied to the 2D material to place it on the substrate surface due to the lattice-mismatch between the two lattices. For the 2D-substrate heterostructure interaction to be attractive, the $\Delta E_{\mathrm{b}} > 0$. In addition, this ΔE_{b} should be greater than the $\Delta E^{f}_{\mathrm{vac}}$ of 2D materials to ensure that the 2D materials remain in their 2D form on the substrate. Singh, Mathew, et al. (2015) has shown previously that the successful synthesis of a 2D material on a particular substrate surface is feasible when the adsorption formation energy, $\Delta E^{f}_{\mathrm{ads}} = \Delta E^{f}_{\mathrm{vac}}$ - $\Delta E_{\mathrm{b}} < 0$.

The framework of *Hetero2d* focuses solely on thermodynamic factors to identify substrates which stabilize meta-stable 2D materials. When the 2D-substrate heterostructure has an $\Delta E^{f}_{\mathrm{ads}} < 0$, the 2D material is stable while $\Delta E^{f}_{\mathrm{ads}} > 0$ indicates the 2D material is unstable. This process enables one to screen for substrates that stabilize meta-stable 2D materials. These 2D-substrate combinations can then be further characterized to address essential questions regarding the adsorbed 2D material's electronic properties and provide insight into bonding strength, charge transfer, and the transferability of the 2D material.

5.3 Hetero2d: The High-Throughput Implementation of the DFT Approach

5.3.1 Introduction

The *Hetero2d* package is an all-in-one workflow approach to model the heterostructures formed by the arbitrary combinations of 2D materials and substrate surfaces. *Hetero2d* can calculate the ΔE^f_{vac}, ΔE_{b}, and ΔE^f_{ads} for each 2D-substrate heterostructure and store the relevant simulation parameters and post-processing in a queryable MongoDB database that can be interfaced to and accessed by an application programming interface (API) or a web-portal. *Hetero2d* is written in Python 3.6, a high-level coding language widely used on modern scientific computing resources. *Hetero2d* utilizes *MPInterfaces* (Mathew et al. 2016) routines and the robust high-throughput computational tools developed by the Materials Project (Mathew et al. 2017; Jain et al. 2013; Jain et al. 2015; Ong et al. 2013) (MP), namely *atomate*, *FireWorks*, *pymatgen*, and *custodian*. Other automatic workflow frameworks and structure manipulation packages exist such as JARVIS (Choudhary et al. 2020), Atomistic Simulation Recipes (Haastrup et al. 2018; Gjerding et al. 2021), AFLOW (Curtarolo et al. 2012), and qmpy (Kirklin et al. 2015).

Hetero2d's framework is inspired by *atomate*'s straightforward statement-based workflow design to perform complex materials science computations with pre-built workflows that automate various types of DFT calculations. Figure 5.1 illustrates the framework of the *Hetero2d* workflow package. *Hetero2d* extends some powerful high-throughput techniques available in existing community packages and combines them with new routines created for this work to generate 2D-substrate heterostructures, perform vdW-corrected DFT calculations, store the stability related data within

a queryable database, and analyze key properties of the heterostructure. In the following sections, each step–outlined in Figure 5.1–is discussed underscoring the new computational tools developed for *Hetero2d*.

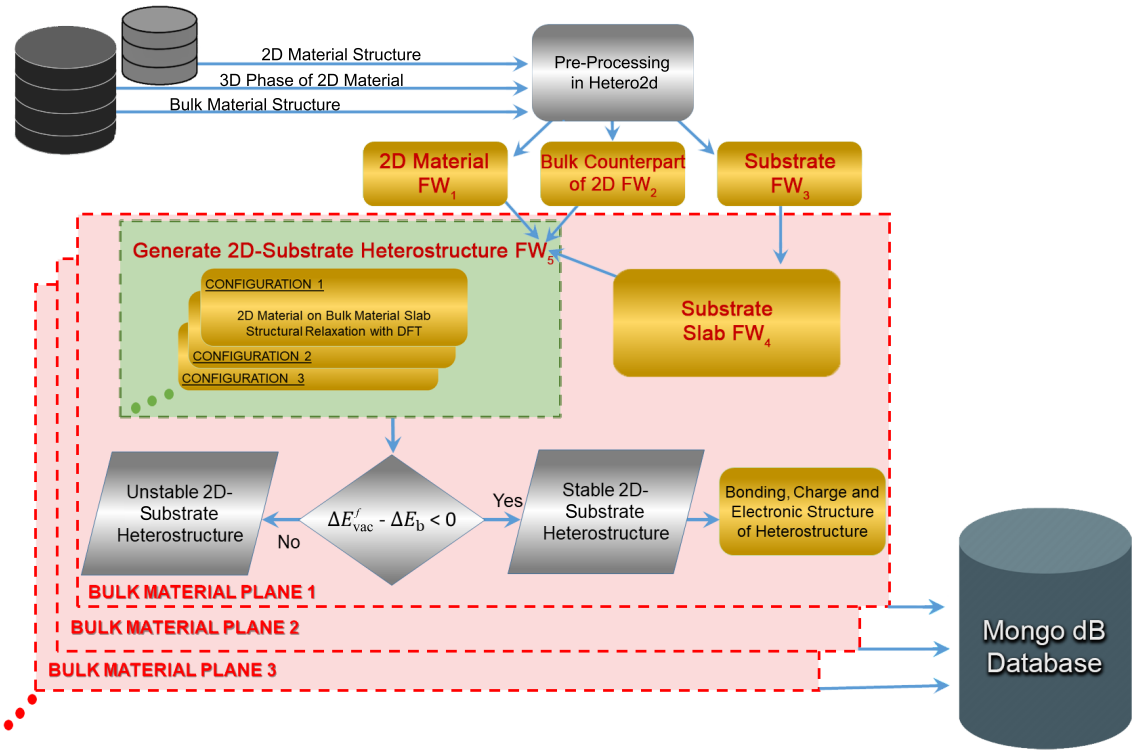

Figure 5.1. Outline of the computational workflow used in this study to investigate the properties of the 2D-substrate heterostructures as coded in the *Hetero2d* package. All structures imported from an external database are relaxed using vdW-corrected DFT with the parameters (discussed below) to maintain consistency. Boxes in gold denote a DFT simulation step and boxes in silver denote a pre-processing or post-processing step.

5.3.2 Workflow Framework

Hetero2d's *atomate*-inspired framework utilizes the *FireWorks* package to break down and organize each task within a workflow. Workflows within the *FireWorks*

package are organized into three task levels–(1) workflow, (2) firework, and (3) firetask. A workflow is a set of fireworks with dependencies and information shared between them through the use of a unique specification file that determines the order of execution of each firework (FW) and firetask. Each FW is composed of one or more related firetasks designed to accomplish a specific task such as DFT structure relaxation. Firetasks are the lowest level task in the workflow. Firetasks can be simple tasks such as writing files, copying files from a previous directory, or more complex tasks such as calling script-based functions to generate 2D-substrate heterostructures, starting and monitoring a DFT calculation, or post-processing a DFT calculation and updating the database.

Hetero2d's workflow *get_heterostructures_stabilityWF* shown in Figure 5.1, has a total of five firework steps–(1) FW_1: the DFT structural optimization of the 2D material, (2) FW_2: the DFT structural optimization of the bulk counterpart of the 2D material, (3) FW_3: the DFT structural optimization of the substrate, (4) FW_4: the creation and DFT structural optimization of the substrate slab, and (5) FW_5: the generation and DFT structural optimization of the 2D-substrate heterostructure configurations. Each firework can be composed of a single or many related firetasks. The tasks are gathered from the specification file that controls the execution of each firetask. For example, FW_1 is used to perform a vdW-corrected DFT structure optimization of the 2D material. Note that the DFT simulations are performed using the Vienna *ab initio* simulation package (Kresse and Joubert 1999; Kresse and Furthmüller 1996b; Kresse and Hafner 1993, 1994; Kresse and Furthmüller 1996a). FW_1 is composed of firetasks which: (1) write VASP input files to the job's launch directory, (2) write the structure file, (3) run VASP using *custodian* (Ong et al. 2013) to perform just-in-time job management, error checking, and error recovery, (4) collect

information regarding the location of the calculation and update the specification file, and (5) perform analysis and convergence checks for the calculation and store all pre-defined information about the calculation in the MongoDB database. A more detailed explanation of each firework in the workflow is discussed in Section 5.3.6, *Workflow Steps*.

5.3.3 Package Functionalities

As mentioned earlier, *Hetero2d* adapts and extends existing community packages to assess the stability of 2D-substrate heterostructures. Table 5.1 lists the functionalities of *Hetero2d* compared with two other workflow-based packages, *MPInterfaces* (Mathew et al. 2016) and *atomate* (Mathew et al. 2017), highlighting new and common features within the three packages.

	Hetero2d	*MPInterfaces*	*Atomate*
Structure processing	✓	✓	✓
Error recovery	✓	✓	✓
Database integration	✓	✓	✓
FireWorks compatible	✓		✓
2D hetero. routines	✓	✓	
2D hetero. workflow	✓		
2D post-processing	✓		

Table 5.1. A list of functionalities present in the *Hetero2d* package compared with two other workflow-based packages *MPInterfaces* and *atomate*. *Hetero2d* is the only workflow package with all the specific features needed to create 2D-substrate heterostructures using high-throughput computational methods.

All three packages utilize the *pymatgen* package to perform various structure processing tasks. *Pymatgen* is used to perform various types of structure-manipulation

processes such as reducing/increasing simulation cell size, creating a vacuum, or creating a slab during the execution of the workflow. Throughout *Hetero2d*, *pymatgen* is utilized to handle structure-manipulation for (a) the bulk materials and (b) some basic pre-/post-processing of structures and generation of files for the DFT calculations. Within *Hetero2d*, *pymatgen*'s structure-manipulation tools are used to create conventional unit cells for the substrate and create the substrate slab surface. Additionally, *pymatgen*'s structure analysis modules have been integrated within *Hetero2d* to decorate the fireworks in the workflow with structural information for each input structure to populate the database. The pre-processing enables one to differentiate crystal phases with similar compound formulas, easily reference and sort data within the database, and perform analysis in later fireworks.

All three packages use the *custodian* package (Ong et al. 2013) to perform error recovery. Error recovery routines are pivotal for any workflow package to reduce the need for human intervention and correct simple run-time errors with pre-defined functions. Additionally, *custodian* alerts the user if an unrecoverable error has occurred.

Database integration is another functionality present in all three packages that stores and analyzes the vast amount of information generated by each calculation.

Only *Hetero2d* and *atomate* are *FireWorks* compatible, whereas *MPInterfaces* uses the python package *fabric* to remote launch jobs over SSH. *FireWorks* is a single package used to define, manage, and execute scientific workflows with built-in failure-detection routines capable of concurrent job execution and remote job tracking over an arbitrary number of computing resources accessible from a clean and flexible Python API.

Routines used to automate the generation of 2D-substrate heterostructures given user constraints are available in *Hetero2d* and *MPInterfaces*. *MPInterfaces* implements

a mathematical algorithm developed by Zur and McGill (1984) for generating supercells of lattice-matched heterostructures given two arbitrary lattices and user-specified tolerances for the lattice-mismatch and heterostructure surface area. *Hetero2d* incorporates functions from *MPInterfaces* to create 2D-substrate heterostructures and enable the *Hetero2d* package to utilize *FireWorks* which *MPInterfaces* is currently incompatible with. Additionally, by incorporating these routines in *Hetero2d*, the code-base can be modify to return and document critical information regarding the 2D-substrate heterostructures that is not returned by *MPInterfaces*. The 2D-substrate heterostructure function returns the strain of the 2D material along **a** and **b** lattice vectors, angle mismatch between the **ab** lattice vectors of the substrate and the 2D material, and scaling matrix used to generate the aligned 2D-substrate heterostructures.

The 2D-substrate heterostructure workflow and post-processing routines are uniquely available in *Hetero2d*. The workflow automates all steps needed to study 2D-substrate heterostructure stability and properties via the DFT method. The post-processing routines enable a curated database to view all calculation results and perform additional analysis or calculations.

5.3.4 Default Computational Parameters

CMDLInterfaceSet is based on *pymatgen*'s *VASPInputSet* class that creates custom input files for DFT calculations. The *CMDLInterfaceSet* class has all the functionality of the parent *pymatgen* class but tailored to perform structural optimizations of 2D-substrate heterostructures and implements vdW-corrections, on-the-fly dipole corrections for slabs, generation of custom k-point mesh grid density, and addition of selective dynamics tags for the 2D-substrate structures. All DFT calculations

are performed using the projector-augmented wave method as implemented in the plane-wave code VASP (Kresse and Joubert 1999; Kresse and Furthmüller 1996b; Kresse and Hafner 1993, 1994; Kresse and Furthmüller 1996a). The vdW interactions between the 2D material and substrate are modeled using the vdW–DF (Rydberg et al. 2003) functional with the optB88 exchange functional (Bowler and Michaelides 2011).

The *CMDLInterfaceSet* has a default energy cutoff of 520 eV used for all calculations to ensure consistency between structures that have the cell shape and volume relaxed, and those that only have ionic positions relaxed. The default k-point grid density was automated using *pymatgen* (Ong et al. 2013) routines to 20 k-points/unit length by taking the nearest integer value after multiplying $\frac{1}{\mathbf{a}}$ and $\frac{1}{\mathbf{b}}$ by 20. These settings were sufficient to converge all calculations to a total force per atom of less than 0.02 eV/Å. Additional information regarding default settings set in the *CMDLInterfaceSet* and convergence tests performed to benchmark the calculations are in the Section C.1 and Section C.2 of Appendix C.

5.3.5 Workflow Initialization and Customization

To use *Hetero2d*'s workflow, *get_heterostructures_stabilityWF*, the 2D structure, its bulk counterpart, and the substrate structure are imported from existing databases through their respective APIs. When initialized, the workflow can accept up to three structures: (1) the 2D structure, (2) the bulk counterpart of the 2D structure, and (3) the substrate structure in the bulk or slab form.

To perform structure transformations to generate the substrate slabs or the 2D-substrate heterostructures, the workflow requires two dictionaries during initialization–

the (1) *h_params* and (2) *slab_params* dictionary. Figure 5.2 is a code excerpt demonstrating the parameters one can supply to generate a 2D-substrate heterostructure on a (111) substrate slab surface. Additional details regarding workflow customization options and current functionality available in *Hetero2d* are discussed in Appendix C Section C.3 and available on CMDLab GitHub as well as an example jupyter notebook.

```
from hetero2d.workflow import  get_heterostructures_stabilityWF

# set substrate slab & 2D hetero params
slab_params = {'transformations': ['SlabTransformation'], 'transformation_params':
   [{'miller_index': [1,1,1], 'min_vacuum_size': 19, 'min_slab_size': 12 }]}
h_params = [{'max_mismatch': 0.05, 'max_area ': 130, 'nlayers_2d': 3, 'nlayers_sub': 2,
   'separation': 3.0 }]
wf = get_heterostructures_stabilityWF(struct_2d, struct_bulk, struct_3d2d, h_params, slab_params)
```

Figure 5.2. Simplified workflow illustrating the setup necessary to setup the 2D-substrate heterostructure workflows using *get_heterostructures_stabilityWF* used throughout this work. A full example jupyter notebook can be found on the CMDLab GitHub.

5.3.6 Workflow Steps

As mentioned previously, the workflow has five firework steps. Here, each preprocessing step involved in initializing the workflow is discussed as well as each firework and their associated firetask(s) for the 2D-substrate heterostructure workflow introduced in Section 5.3.2, *Workflow Framework*.

The first firework, FW_1, in the workflow optimizes the 2D material structure. During initialization of the workflow, the 2D material is centered within the simulation cell, obtaining crystallographic information regarding the structure, the *CMDLInterfaceSet* is initialized to create VASP input files, and a list of user-defined/default tags

are created for the 2D material. The structure, tags, and *CMDLInterfaceSet* are used to initialize the firework *HeteroOptimizeFW* that performs the structure optimization. The default tags appended to the firework are the unique identification tags (provided to the workflow by the user), the crystallographic information, workflow and firework name, and the structure's composition. In FW_1, *HeteroOptimizeFW* executes firetasks that–(a) create directories for the firework, (b) write all input files initialized using *CMDLInterfaceSet*, (c) submit the VASP calculation to supercomputing resources to perform full structure optimization and monitor the calculation to correct errors, (d) run the *HeteroAnalysisToDb* class to store all information necessary for data analysis within the database, and (e) lastly pass the information to the next firework. Details regarding *HeteroAnalysisToDb* can be found in the next section.

Similar to FW_1, FW_2, and FW_3 perform a full structural optimization for the bulk counterpart of the 2D material and the substrate, respectively. FW_2 and FW_3 differ from FW_1 only in the pre-processing steps. The step to center the 2D material is not performed; however, the conventional standard structure is utilized during the pre-processing for FW_3.

FW_3 spawns a child firework passing the optimized substrate structure to FW_4 which transforms the conventional unit cell of the substrate into a substrate slab using the *slab_params* dictionary and performs the structure optimization. When the workflow is initialized, FW_4 undergoes similar pre-processing steps that are used to initialize the firework *SubstrateSlabFW* that creates a substrate slab from the substrate. *SubstrateSlabFW* is the firework that transforms the conventional unit cell of the substrate into a slab, sets the selective dynamics tags on the surface layers, and sets the number of compute nodes necessary to relax the substrate slab. The *slab_params* variable is the input dictionary that initializes *pymatgen*'s *SlabTransformation* module

that creates the substrate slab. All required and optional input arguments used in the *SlabTransformation* module must be supplied using this dictionary (key: value) format. This dictionary format is implemented to enable *Hetero2d* to be flexible and extendable in future updates. Additionally, the *slab_params* dictionary is only required when creating a new substrate slab from a substrate.

After the first four fireworks have been completed and successfully stored in the database, the fifth firework (FW_5) obtains the optimized structures and information from previous fireworks and the specification file. FW_5 calls the *GenHeteroStructuresFW* firework to generate the 2D-substrate heterostructure configurations using *h_params* and spawns a firework to perform structure optimization for each configuration. The input required for the *h_params* dictionary are those that are required by *Hetero2d*'s *hetero_interfaces* function. This function attempts to find a matching lattice between the substrate surface and the 2D material. The parameters used to initialize *hetero_interfaces* are listed in the *h_params* dictionary shown in Figure 5.2.

The function *hetero_interfaces* generates the 2D-substrate heterostructure configurations utilizing *MPInterfaces*'s interface matching algorithm. To ensure functions within the workflow are compatible with *FireWorks* the *hetero_interfaces* function was adapted from *MPInterfaces*. Additionally, this ensures key variables regarding the interfacing matching algorithm–such as the strain or angle mismatch–are returned and stored in the database. *MPInterfaces* is used to (a) generate heterostructures within an allowed lattice-mismatch and surface area of the supercell at any rotation between the 2D material and bulk material surface and (b) create distinct configurations in which the 2D material can be placed on the bulk material surface based on the Wyckoff positions of the near-interface atoms.

FW_5 calls *GenHeteroStructuresFW* which generates the 2D-substrate heterostruc-

ture configurations, the total number of configurations is computed, each unique configuration is labeled from 0 to n-1, where n is the total number of configurations, and stored under the *Interface Config* tag. For each configuration, a new firework is spawned to optimize each 2D-substrate heterostructure configuration. The data generated within FW_5 is stored in the database.

After all previous FWs have successfully converged, *HeteroAnalysisToDb* is called one final time to compute the ΔE^f_{vac}, ΔE_{b}, and ΔE^f_{ads} for each heterostructure configuration generated by the workflow. The calculation of the ΔE^f_{vac} references the simulation for the 2D material and its bulk counterpart. The bulk counterpart is simulated using a standard periodic simulation cell. The calculation of ΔE_{b} references the 2D material, substrate slab, and 2D-substrate heterostructure simulations which all employ a standard supercell slab model. The calculation of the ΔE^f_{ads} references both ΔE_{b} and ΔE^f_{vac}. Once each value is computed, all the information is curated and stored in the MongoDB database.

5.3.7 Post-Processing Throughout the Workflow

After each VASP simulation is complete, post-processing is performed within the calculation directory using the *HeteroAnalysisToDb* class, an adaptation of *atomate*'s *VaspToDb* module. It is used to parse the calculation directory, perform error checks, and curate a wide range of input parameters and quantities from calculation parameters and output, energetic parameters, and structural information for storage in the MongoDB. *HeteroAnalysisToDb* detects the type of calculation performed within the workflow and parses the calculation accordingly. *HeteroAnalysisToDb* has the same functionally as *VaspToDb* with additional analyzers developed for 2D-substrate

heterostructures that–(a) identify layer-by-layer interface atom IDs for the substrate and 2D material, (b) store the initial and final configuration of all structures, (c) compute the $\Delta E^{f}_{\mathrm{vac}}$, ΔE_{b}, and $\Delta E^{f}_{\mathrm{ads}}$, (d) store the results obtained from the interface matching, and (e) ensure each database entry has any custom tags added to the database such as those appended by the user. The workflow design ensures that the DFT simulations for each 2D-substrate surface pair will be performed independently of each other, but as soon as all simulations are completed for each 2D-substrate surface pair, the data will be analyzed and curated in the MongoDB database right away.

5.4 An Example of Substrate Screening via Hetero2d

5.4.1 Materials Selection

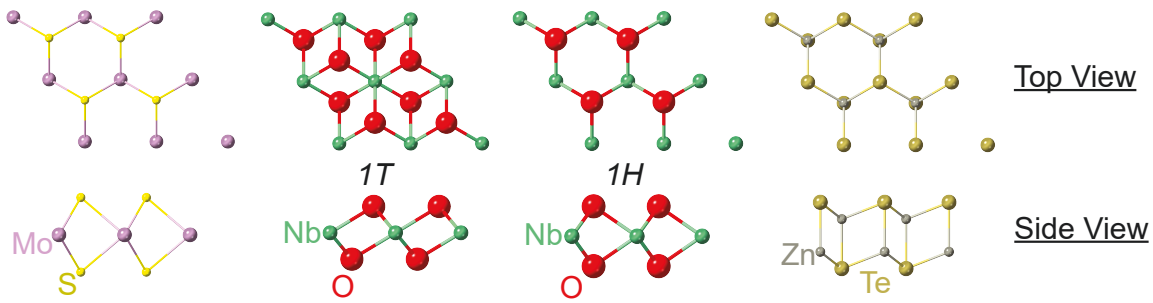

Figure 5.3. Structure models illustrating the 2D films crystal structure. Top view demonstrates the hexagonal symmetry of each 2D material. The $1T$ and $1H$ phase for NbO_2 are labeled to clarify the two phases.

To demonstrate the functionalities of the *Hetero2d* package, four 2D materials–namely $1H$-MoS_2, $1T$-NbO_2, $1H$-NbO_2 (Haastrup et al. 2018), and hexagonal-ZnTe (S.

Torrisi et al. 2020)–are screened for suitable substrate matches. The four 2D materials in consideration possess hexagonal symmetry as illustrated in Figure 5.3.

MoS_2 was selected because there is a large amount of experimental and computational (Chen et al. 2013; Zhuang and Hennig 2013; Yun et al. 2012; Singh, Hennig, et al. 2015) data available in literature which are used to validate the computed properties from using the *Hetero2d* workflow. The hexagonal-ZnTe (S. Torrisi et al. 2020), 1T-NbO_2, and 1H-NbO_2 (Haastrup et al. 2018) are yet to be synthesized. In addition, these particular 2D materials have diverse predicted properties; see Table 5.2. It is noteworthy that hexagonal-ZnTe has been predicted to be an excellent CO_2 reduction photocatalyst (S. Torrisi et al. 2020).

2D Mat.	MoS_2	1T-NbO_2	1H-NbO_2	ZnTe
Classification	Semiconductor	FM [1.]	FM [1.]	Semiconductor
Band Gap (eV)	1.88 [2.]	0.0 [1.]	0.0 [1.]	2.88 [3.]

Table 5.2. The electronic properties and band gap of the four selected 2D materials used in this work. FM represents ferromagnetic. Works cited 1. Haastrup et al. (2018) 2. Gusakova et al. (2017) 3. S. Torrisi et al. (2020).

The properties of a 2D material can differ when placed on different miller-index planes for the same substrate. Thus, all unique low-index substrate surfaces (with h, k, l equal to 1 or 0) for these 2D materials are investigated. A material available in the Materials Project (MP) (Ong et al. 2013) database was considered a potential substrate if it satisfied all of the following criteria–(a) is metallic, (b) is a cubic phase, (c) is single-element composition, (d) has a valid ICSD ID (Standards and Technology) (thus been experimentally synthesized), and (e) has an $E_{above\ hull} < 0.1$ eV/atom. There are 50 total substrates that satisfy the criteria above when queried from the MP database.

The bulk counterpart of each 2D material is also obtained from the MP database. The database is query-ed for bulk materials that have the same composition as the 2D material and select the structure with the lowest $E_{above\ hull}$. Appendix C Table C.3 - Table C.2 have additional reference information regarding all the optimized substrate slabs, 2D materials, and their bulk counterparts. Appendix C Table C.3 contains information about the Materials Project material_id, $E_{above\ hull}$, ICSD ID, crystal system, and miller plane for the substrate surface. Appendix C Table C.1 contains information about the reference database ID, ΔE^{f}_{vac} (eV/atom), and crystal system for each 2D material and Appendix C Table C.2 contains information about the reference database id, $E_{above\ hull}$, E_{gap}, and the crystal system for the bulk counterpart of the 2D material.

5.4.2 Symmetry-Matched, Lattice-Matched 2D-Substrate Heterostructures

This study focuses the search for 2D-substrate heterostructures to substrate planes with indices, h, k, l as 0 or 1. The following studies focus on the heterostructures with the (111) and (110) substrate surfaces since these two miller planes have an appreciable number of heterostructures. The (001) substrate plane resulted in only one heterostructure.

2D Mat.	(111) Substrate	(110) Substrate
MoS_2	Hf, Ir, Pd, Zr, Re, Rh	Ta, Rh, Sc, Pb, W, Y
1T-NbO_2	Ni, Mn, V, Nd, Pd, Ir, Hf, Zr, Cu	Rh, Ta, Sc, W
1H-NbO_2	Ni, Mn, Nd, Ir, Hf, Al, Te, Ag, Ti, Cu, Au	Ta, Sc, W, Y, Rh
ZnTe	Sr, Ni, Mn, V, Al, Ti, Cu	W

Table 5.3. A list of matching substrate surfaces for the four 2D materials given the heterostructure search criteria discussed in the next section.

Restricting the search for 2D-substrate matches to only the (111) and (110) yields a total of 4 (# of 2D materials) X 2 (# of planes) X 50 (# of substrates) = 400 potential 2D-substrate heterostructure combinations. As illustrated in Figure 5.4, after introducing the constraints for the surface area (σ_{SA}) < 130 Å^2 and applied strain on the 2D material to be < 5%, a total of 49 2D-substrate heterostructure workflows are found. Table 5.3 lists all metallic substrates matching each of the 2D materials given the heterostructure criteria.

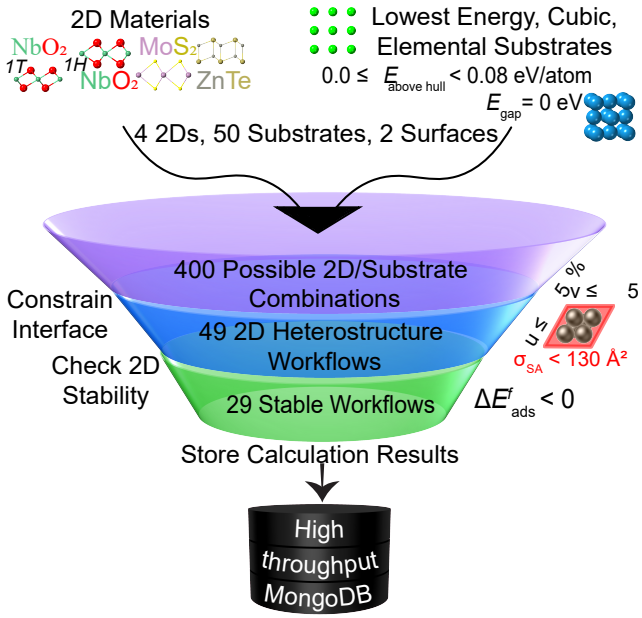

Figure 5.4. Schematic representing the materials selection process identifying stable 2D-substrate heterostructures using the *Hetero2d* workflow. Tier 1 represents choosing 2D materials, substrates, and their surfaces. Tier 2 applies constraints on the surface area and lattice strain. Tier 3 shows the energetic stability of the heterostructures stored in the database.

Of the total 49 workflows, 33 workflows correspond to the (111) substrate surfaces, and 16 workflows correspond to the (110) substrate surfaces. Generally, the (111) surface has more substrate matches than (110) surface due to the intrinsic hexagonal symmetry of the (111) surface that matches the hexagonal symmetry of the selected

2D materials. Each workflow generates between 2-4 2D-substrate heterostructure configurations for a given 2D-substrate surface pair, resulting in a total of 123 2D-substrate heterostructure configurations. Of those 2D-substrate heterostructures, 78 configurations, a total of 29 workflows stabilize the meta-stable 2D materials when placed upon the substrate slab. Additional details regarding these simulations can be found in Appendix C Section C.4.

5.4.3 Stability of Free-Standing and Adsorbed 2D Films and Heterostructures

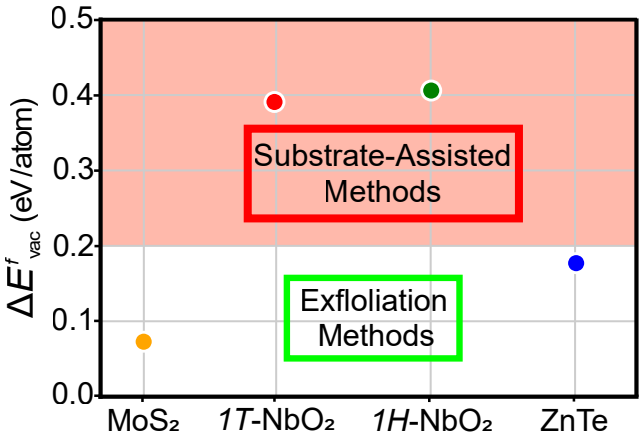

Figure 5.5. The ΔE^f_{vac} for 2D MoS_2 (•), 1T-NbO_2 (•), 1H-NbO_2 (•), and ZnTe (•). The ΔE^f_{vac} is used to assess the thermodynamic stability of the free-standing 2D film with respect to its bulk counterpart. MoS_2 and ZnTe have relatively low ΔE^f_{vac} while the 1T and 1H phase of NbO_2 have high ΔE^f_{vac}.

Figure 5.5 shows the ΔE^f_{vac} of the isolated unstrained 2D material with respect to their bulk counterpart. The ΔE^f_{vac} for both MoS_2 and ZnTe are low, less than 0.2 eV/atom. Both the 1T and 1H phase for NbO_2 possess high ΔE^f_{vac}, as shown by the red shaded region in Figure 5.5, making substrate-assisted synthesis methods the most feasible method to synthesize these 2D films. The ΔE^f_{vac}'s in Figure 5.5 are

consistent with prior computational (Haastrup et al. 2018; S. Torrisi et al. 2020) and experimental work (Y.-H. Lee et al. 2013).

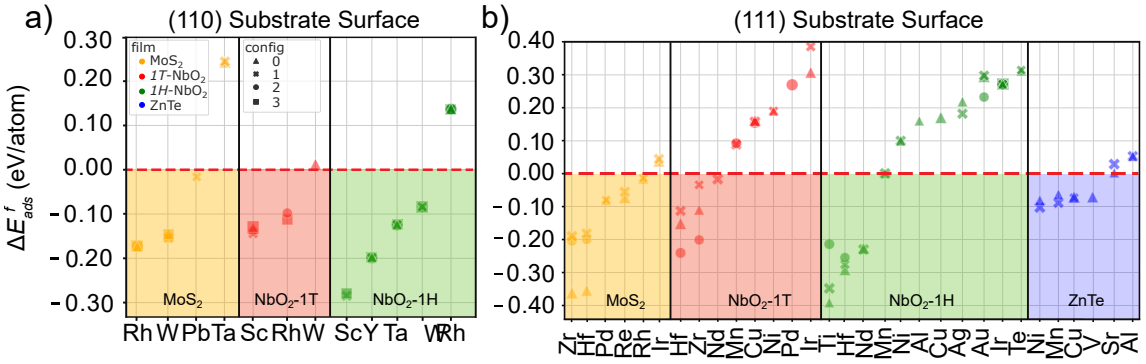

Figure 5.6. Adsorption formation energy, ΔE^f_{ads}, for the symmetry-matched, low lattice-mismatched (a) (110) and (b) (111) substrate surfaces. The rectangular symmetry of the (110) surface results in fewer matches while the hexagonal symmetry of the (111) substrate surface results in numerous matches within the given constraints on the surface area and lattice strain. Negative ΔE^f_{ads} values indicate stabilization of the 2D material. Each set of symbols (up to 4 points per substrate) represents the unique 2D-substrate configurations.

Figures 5.6a and Figure 5.6b show the ΔE^f_{ads} for the four 2D materials on the (110) and (111) substrate surfaces, respectively. The vertical black lines in Figure 5.6 separate the 2D materials, while the shaded regions indicate stabilization of the 2D material on the substrate surface. When generating 2D-substrate heterostructure, the first challenge is finding a matching lattice between the 2D material and substrate surface. The next challenge is identifying "ideal" or likely locations to place the 2D material on the substrate surface to generate stable low-energy heterostructures. To reduce the large number of in-plane shifts possible for a given 2D-substrate heterostructure, the 2D material is selectively placed on the substrate slab enumerating combinations of high-symmetry points (Wyckoff sites) between the 2D material and substrate slab stacking the 2D material on top of these sites z Å away from the substrate surface.

Each unique 2D-substrate heterostructure configuration is represented by 0=△, 1=**x**, 2=∘, and 3=□ in Figure 5.6.

The ΔE^f_{ads} on the (110) surface is shown in Figure 5.6a. There are 9 substrates which stabilized the ΔE^f_{ads} of the 2D materials. The ΔE^f_{ads} appears to be correlated with the substrate where the 2D material is placed; however, the data is to sparse in Figure 5.6a to distinguish the origin of this trend. Interestingly, when MoS_2 is placed on the (110) Ta substrate surface, the 2D material buckles which likely increases the ΔE^f_{ads} significantly above the other substrates. Appendix C Figure C.4 shows both configurations for MoS_2 on the (110) Ta substrate surface. There are an additional 5 2D-(110) substrate pairs that were studied but are not shown in Figure 5.6a because the 2D materials/substrate interface becomes highly distorted/completely disintegrated. These cases are shown in Appendix Figure C.2a and discussed in Section C.5.

The (111) substrate surface matches for each 2D material are shown in Figure 5.6b, where 15 substrates result in an $\Delta E^f_{\text{ads}} < 0$. An additional 8 2D-substrate pairs, shown in Appendix Figure C.2b, have 2D materials/substrate surfaces that are disintegrated and are discussed in Section C.5 in Appendix C.

A correlation between the substrate surface and the ΔE^f_{ads} is more apparent for the (111) surface in Figure 5.6b due to the increased number of 2D-substrate pairs. For MoS_2 on Zr and Hf, the triangle configurations have ΔE^f_{ads} significantly lower than the other configurations, see Appendix C Figure C.4 for structures of the three configurations. The lower ΔE^f_{ads} is correlated with smaller bond distances between the substrate surface and the 2D material. For these structures, the ΔE^f_{ads} is found to be lower for structures where the $2h$ Wyckoff site of the 2D material is stacked on top of the $2a$ Wyckoff site of the substrate surface. The location of a 2D material on a substrate surface has previously been shown to influence the type of bonding

present between the 2D material and substrate surface (Singh, Zhuang, and Hennig 2014; Zhuang et al. 2017).

The 1T phase of NbO_2 on Hf, Zr, and Ir substrates have an ΔE^f_{ads} difference between each configuration that is larger than other 2D-substrate pairs. The differences in ΔE^f_{ads} for 1T-NbO_2 on Ir is partly due to some structural disorder of the 2D materials from the O atoms bonding strongly with the substrate surface, shown in Appendix C Figure C.5. For both Hf and Zr, the differences in ΔE^f_{ads} do not arise from structural disorder. The ΔE^f_{ads} of 1T-NbO_2 on Hf and Zr are more strongly affected by the location of the 2D material on the substrate surface.

1H-NbO_2 has two substrate surfaces, Ti and Au, where the ΔE^f_{ads} varies strongly with the configuration of 2D material on the substrate, unlike other 2D-substrate pairs for 1H-NbO_2. 1H-NbO_2 on Ti and Au have no structural distortions that explain the difference in ΔE^f_{ads}. For 1H-NbO_2 on Ti, each configuration possesses different ΔE^f_{ads} arising from the unique placement of the 2D material on the substrate surface for each configuration. The strong bonding between the 2D material and substrate surface may be due to the affinity for Ti to form a metal oxide. Appendix C Figure C.6 shows each configuration for 1H-NbO_2 on (111) Ti substrate surface. For 1H-NbO_2 on Au, the circle configuration has a lower ΔE^f_{ads} due to the bottom layer of the 1H-NbO_2 stacked directly on the top layer of the Au substrate surface.

The properties of MoS_2 have been studied both computationally and experimentally, where previous computational works (Zhuang and Hennig 2013; Singh, Mathew, et al. 2015) have found similar values for the ΔE^f_{vac} of MoS_2. Chen et al. (2013) found that Ir bonds more strongly with the substrate surface than Pd. This may explain the small structural modulations observed in this study for MoS_2 on the Ir (111) substrate surface but no such modulation is observed for MoS_2 on the Pd (111) substrate surface.

Additionally, the z-separation distance between the 2D material and substrate surface found in this work agrees well with (Chen et al. 2013)'s values despite using a different functional. The z-separation distances are within 0.05 Å for Ir and 0.16 Å for Pd.

In summary, MoS_2 and ZnTe have low formation energies and are potentially well suited for exfoliation-based methods (Singh, Mathew, et al. 2015; Revard et al. 2016) while the $1T$- and $1H$-NbO_2 phase have high formation energies requiring substrate assisted methods to synthesize these metastable 2D materials. Through the substrate screening process, a total of 9 (110) and 15 (111) substrate slab surfaces are found to stabilize the 2D materials considered here. MoS_2 is stabilized by the (110) surfaces of Rh, W, and Pb and (111) surfaces of Zr, Hf, Pd, Re, and Rh. The $1T$-NbO_2 is found to be stabilized by the (110) surfaces of Sc, and Rh and (111) surfaces of Hf, Zr, and Nd. The $1H$-NbO_2 is found to be stabilized by the (110) surfaces of Sc, Y, Ta, and W and (111) surfaces of Ti, Hf, and Nd. The ZnTe is stabilized only by the (111) surfaces of Ni, Mn, Cu, and V.

5.4.4 Thickness Changes in Adsorbed 2D Films on Substrate Slab Surfaces

The change in the thickness of the adsorbed 2D material may provide insight into the nature of bonding between the 2D-substrate heterostructures. For instance, vdW bonds are weak and thus typically result in minimal structural and electronic changes in the 2D material. Using the database, the change in the thickness of post-adsorbed 2D materials from that of the free-standing 2D material is computed. The thickness of the free-standing/adsorbed 2D material is computed first by finding the average z coordinate of the top and bottom layer of the 2D material given by $\bar{d}_z = \sum_{i=1}^{n} d_{i,z}^{top}/n - \sum_{i=1}^{m} d_{i,z}^{bottom}/m$ where $d_{i,z}$ is the z coordinate of the i^{th} atom summed

up to n and m, the total number of atoms in the top and bottom layers, respectively. The thickness, obtained by taking the difference between the average thickness of the adsorbed 2D material from that of the free-standing 2D material, $\delta d = \bar{d}_z^{adsorbed} - \bar{d}_z^{free}$, with positive (negative) values corresponding to an increase (decrease) in the thickness of the adsorbed 2D material.

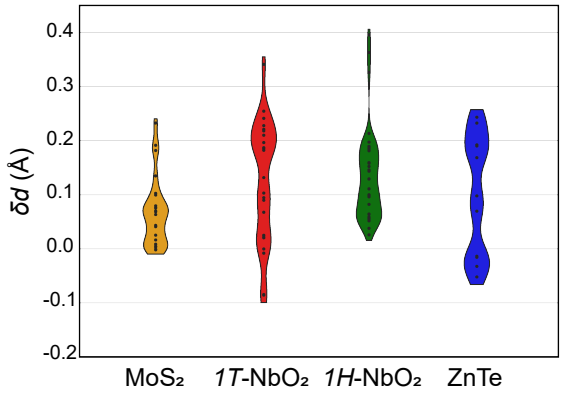

Figure 5.7. Each 2D material is separated spatially along the x-axis using a violin plot. The change in the 2D material's thickness, δd, for all substrates is plotted along the y-axis. A positive y-value indicates the 2D material's thickness has increased during adsorption onto the substrate slab. The width of the violin plot is non-quantitative from scaling the density curve by the number of counts per violin; however, within one violin plot, the relative x-width does represent the frequency that a 2D material's thickness changes by y amount relative to the total number of data points in the plot.

Figure 5.7 illustrates the change in the thickness of the free-standing 2D material from that of the adsorbed 2D material for each 2D-substrate heterostructure. Typically for vdW type bonding, each atom should have minimum deviations from the free-standing 2D film due to the weak interaction between the adsorbed 2D material and substrate surface that characterizes vdW bonding. Figure 5.7 shows many of the 2D-substrate pairs have a significant change in the thickness of the 2D material that may indicate more covalent/ionic type bonding. The change in the thickness of the 2D material for the majority of the MoS_2-substrate configurations is minimal (< 0.1

Å) that may indicate weak interactions between the 2D material and substrate surface. Figure 5.7 indicates that for the majority of the adsorbed 2D materials, the substrates tend to increase the thickness of the adsorbed 2D material.

Analysis of the 2D material's thickness shows most adsorbed 2D materials thickness has increased substantially which may indicate the interaction between the 2D material and the substrate surface is covalent/ionic. A characterization of the density of states and the charge transfer (through Bader analysis or charge density difference), as presented in the next section, is required to determine the exact nature of the interaction.

5.4.5 Charge Layer Doping of Adsorbed 2D Films

The *Hetero2d* workflow package has a similar infrastructure as *atomate* that allows the package to integrate seamlessly with the workflows developed within *atomate*. These atomate workflows enable the information generated by *Hetero2d* to be easily expand upon by utilizing existing workflow to perform various analysis such as Bader (Tang, Sanville, and Henkelman 2009; Henkelman, Arnaldsson, and Jønsson 2006) charge analysis and high-quality density of states (DOS) calculations to assess charge transfer that occurs between the adsorbed 2D material and the substrate surface, changes in the DOS from the adsorbed and pristine 2D material, and changes in the charged state of the 2D-substrate pairs.

Most 2D materials are desirable due to their unique electronic properties. MoS_2 on Hf (111) surface was selected to demonstrate the capability of *Hetero2d* in providing detailed electronic and structural information. The Bader analysis illustrated in Table

5.4 shows that there is charge transfer from the substrate to the bottom layer of the 2D material which is consistent with the findings presented by Zhuang et al. (2017).

electrons	Q_{sub}	Q_{S_b}	Q_{Mo}	Q_{S_t}	$Q_{S_b}^{prist}$	Q_{Mo}^{prist}	$Q_{S_t}^{prist}$
Q_x	-0.11	1.10	-1.03	0.57	0.60	-1.20	0.60

Table 5.4. Q_x is obtained with Bader analysis and represents the average number of electrons transferred to/from (positive/negative) specific atomic layers with the initial number of electrons taken from the POTCAR. The first four columns are the electrons transferred to/from–the Hf substrate atoms, Q_{sub}, the bottom layer of S atoms, Q_{S_b}, the Mo atoms, Q_{Mo}, and the top layer of S atoms, Q_{S_t} for the adsorbed 2D-substrate heterostructure. The last three columns denote the charge transfer in the pristine MoS_2 structure. MoS_2 has an increased charge accumulation on the bottom layer of the 2D material from the substrate slab.

In Figure 5.8a, the DOS for the isolated un-strained, isolated strained, and adsorbed MoS_2 is shown where the black dashed line represents the Fermi level. There is a small shift in the DOS when comparing the un-strained and strained DOS for MoS_2. Comparing the DOS for the adsorbed MoS_2 to the other DOS for MoS_2, there is a significant change in the DOS. when MoS_2 is placed on the Hf (111) surface a semiconductor to metal transition is seen in the DOS of MoS_2 induced by the substrate. This change in the DOS is consistent with the Bader analysis that indicates electron doping of the MoS_2 material occurs which would result in changes in the DOS. Figure 5.8b shows the redistribution of charge due to the interaction of the 2D material and substrate surface where red and blue regions indicate charge accumulation (gaining electrons) and depletion (losing electrons) of the combined system due to the interaction between MoS_2 and Hf. The charge density difference is computed as the difference between the sum of the isolated MoS_2 and isolated Hf substrate slab from that of the combined MoS_2 on Hf system. Figure 5.8c is the charge density of the combined MoS_2 on Hf system along the (110) plane. Thus, the electronic

properties of MoS_2 are dramatically affected by the substrate. *Hetero2d* can analyze the substrate induced changes in the electronic structure of 2D materials. This will lead to a fundamental understanding and engineering of complex interfaces.

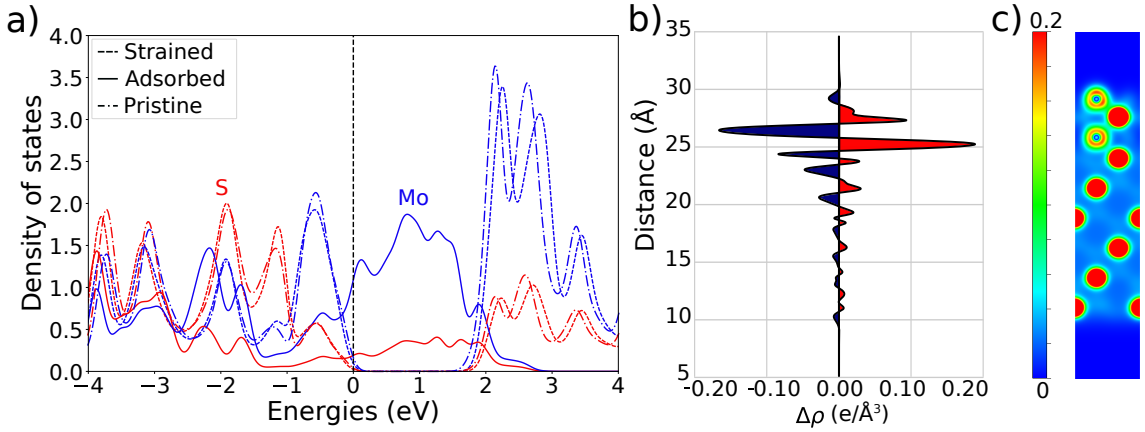

Figure 5.8. (a) The element projected density of states (DOS) where red and blue lines correspond to S and Mo states, respectively, for the isolated strained 2D material (dashed lines), the adsorbed 2D material (solid lines), and the pristine MoS_2 material (dashed-dotted lines). The Hf (111) substrate influences the DOS for MoS_2 causing a semiconductor to metal transition. (b) The z plane-averaged electron density difference ($\Delta\rho$) for MoS_2 on Hf. Electron density difference is computed by summing the charge density for the isolated MoS_2 and isolated Hf then subtracting that from the charge density of the interacting MoS_2 on Hf system. The charge densities were computing with fixed geometries. The red and blue colors indicate electron accumulation and depletion in the combined MoS_2 on Hf system, respectively, compared to the isolated MoS_2 and isolated Hf atoms. (c) The charge density distribution for MoS_2 on (111) Hf substrate. The cross section is taken along the (110) plane passing through Mo, S, and Hf atoms. The charge density is in units of electrons/Å^3.

These results show a semiconductor-to-metal transition for the MoS_2 on Hf which is consistent with previous calculations by Zhuang et al. (2017). Additionally, charge transfer from the Hf substrate to MoS_2 is shown.

5.5 Summary

In summary, this work discusses the development of *Hetero2d*, an open-source workflow package that automates the generation of 2D-substrate heterostructures, the creation of DFT input files, the submission and monitoring of computational jobs on supercomputing facilities, and the storage of relevant parameters alongside the post-processed results in a MongoDB database. Using the example of four candidate 2D materials and low-index planes of 50 potential substrates the utility of *Hetero2d* to address the immense number of 2D material-substrate surface pairs to guide the experimental realization of novel 2D materials is demonstrated. Among the 123 configurations studied, only 78 configurations (29 workflows) are found to result in stable 2D-substrate heterostructures. The use of *Hetero2d* is further illustrated by examining the change in thickness of the adsorbed 2D materials, the Bader charges, and the electronic density of states of the heterostructures to study fundamental changes in the properties of the post-adsorbed 2D material. *Hetero2d* is freely available on the CMDLab GitHub website under the GNU license along with example jupyter notebooks.

These results show the four 2D materials, MoS_2 and ZnTe have low formation energies while both the $1T$- and $1H$-NbO_2 phases have high formation energies requiring a substrate to feasibly synthesize the 2D material. Through the substrate screening process, a total of 9 (110) and 15 (111) substrate surfaces are found to stabilize the 2D materials considered here. An analysis of the 2D material's thickness shows most adsorbed 2D materials thickness has increased substantially, which may indicate the interaction between the 2D material and the substrate surface is covalent/ionic. From the density of states and Bader charge analysis, a semiconductor-to-metal

transition in 2D MoS_2 adsorbed on Hf is observed due to the significant charge transfer between the Hf substrate and MoS_2.

Chapter 6

FUNDAMENTAL FACTORS GOVERNING SUBSTRATE INDUCED STABILIZATION OF JANUS 2D HETEROSTRUCTURES WITH MACHINE LEARNING

6.1 Motivation

Out-of-plane symmetry breaking is highly desired for 2D materials with the emergence of novel properties for widespread device applications in electronics (Riis-Jensen et al. 2019; Lu et al. 2017; Qin et al. 2022; Ma et al. 2022), catalysis (Idrees, Fawad, et al. 2020; Idrees, Din, et al. 2020; G. Zhang et al. 2021), sensing (Lei Zhang et al. 2020), quantum computing (Maghirang et al. 2019; Y. Zhang et al. 2019), energy applications Zhang2020, Guo2019, and biology (Fu et al. 2020). Janus 2D materials–typically 3 atomic layers thick wherein each atomic layer is a distinct element–offer intrinsic out-of-plane symmetry breaking, giving rise to interesting properties–most notably, strong Rashba spin splitting, finite out-of-plane dipole moments, and in-plane piezoelectricity (Riis-Jensen et al. 2019; Lei Zhang et al. 2020). These properties make Janus 2D materials highly tunable and ideal candidate materials for vertically stacked heterostructure device applications. However, experimental realization and utilization of hypothetical 2D materials thus far has been slow. The C2DB (Haastrup et al. 2018) contains roughly 4,000 predicted 2D materials with $\sim$55 being experimentally synthesized. One challenge in realizing 2D materials is the underlying fundamental physics that govern the impact that substrates have on the post-adsorbed properties

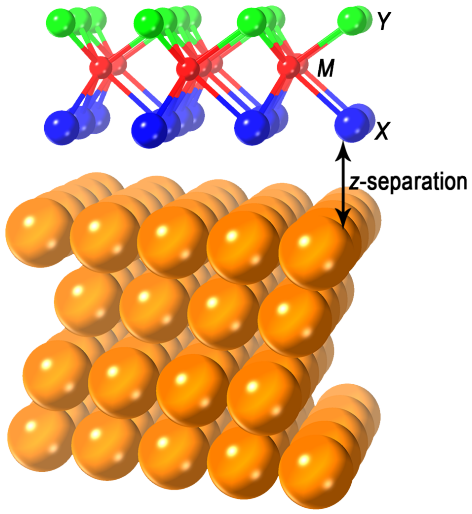

Figure 6.1. 1H-Janus 2D-substrate heterostructure model showing the orientation of the 2D material on top of the substrate surfaces. The Y-M-X represent the atomic layers of the Janus 2D material.

and energetic stability 2D material heterostructures is not well known for the rich physical and chemical space of most predicted 2D materials.

Substrates can play a critical role in determining the 2D material properties. They can serve as a stabilizing agent to metastable phases during growth (Tusche, Meyerheim, and Kirschner 2007) (as a weakly coupled van der Waals (vdW) surface securing the 2D material (Walsh and Hinkle 2017; Qin et al. 2022)) or be used to tune the 2D material properties (Jena and Konar 2007; Singh, Hennig, et al. 2015; S. Xu et al. 2015). Current computational efforts have largely focused on common 2D materials like transition metal dichalcogenides (TMDC), graphene, or hexagonal-BN. The overwhelming majority of predicted 2D materials lack critical information such as substrates that stabilize metastable 2D materials or substrate effects on adsorbed 2D materials. This information is needed to provide guidelines for substrate selection to experimentally synthesize the numerous hypothesized 2D materials as well as their post-adsorbed properties.

In this report, the 2D Materials Synthesis Database (2DMSD) is introduced, the first freely available database of calculated data on 2D-substrate heterostructures and their post-adsorbed properties, all accessible through a web portal and downloadable. Hetero2d (Boland and Singh 2022)–a high-throughput computational synthesis package designed for computing heterostructure properties–is used to study heterostructures of 52 Janus 2D materials (Y=[S, Se, Te], M=[As, Bi, Cr, Hf, Mo, Nb, Sb], and X=[Br, Cl, I, Se, Te], see Figure 6.1) and 19 metallic, cubic phase, elemental substrates on low-index crystal planes for lattice- and symmetry-matched 2D-substrate heterostructures. Figure 6.1 illustrates the vertical stacking of each layer in the Janus 2D material and the substrate surface with the interface forming between the X-layer of the 2D material and top-layer of the substrate. 438 Janus 2D-substrate pairs are analyzed to identify substrate surfaces that stabilize metastable Janus 2D materials, assess substrate effects in post-adsorbed 2D materials, and identify charge transfer between the 2D material and the substrate surface. Additionally, a computational framework combining *ab-initio* calculations with machine learning methods–similar to previous works (Willhelm et al. 2022)–to screen and identify stable and post-adsorbed properties of 2D-bulk substrate heterostructures is utilized. The *ab-initio* calculations are used to generate random forest regression models and predict the binding energy, z-separation distance, and charge transfer to the 2D material for 2D-substrate heterostructures with 0.041 eV/atom, 0.120 Å, and 0.203 e- accuracy (assessed using the root mean squared error), respectively. Additionally, the models elucidate the fundamental factors that lead to the stabilization of 2D materials, the z-separation distance between the substrate surface and 2D material, and the charge transferred to the 2D material.

This work advances the understanding of the fundamental relationships that govern the interactions of 2D materials with substrate surfaces providing rational guidelines

towards the experimental synthesis and utilization of the thousands of predicted 2D materials. The critical property in determining the substrate induced stabilization of a 2D-bulk substrate pair (ΔE_{b}) was found to be the substrate's electronegativity while the critical property in determining the z-separation distance was the substrates surface energy. The primary property to predict the charge transferred to the 2D material was the fraction of p valence electrons. By combining the database and machine learning models, a data-driven approach can be utilized to conduct targeted searches and explore stable 2D-substrate heterostructures and their properties reducing the computational costs typically associated with brute force *ab-initio* based calculations of large combinatorial spaces.

6.2 Screening Parameters and Materials Selection

6.2.1 Materials Selection

Janus 2D materials show promise for quantum computing device applications (Maghirang et al. 2019) originating from their unique out-of-plane inversion symmetry breaking and lack of mirror symmetry due to their distinct X, Y atomic layers. The distinct atomic arrangement creates an out-of-plane intrinsic field from the differing electronegativity which make these materials candidate materials for quantum computing devices among others (Chen et al. 2013; Maghirang et al. 2019; Hou et al. 2022).

The Janus 2D materials were obtained from the C2DB (Haastrup et al. 2018) which contains ∼200 computed Janus 2D materials with a wide range of predicted properties. 52 distinct Janus 2D phases were selected by considering 2D materials

with a corresponding composition matched bulk (3D) counterpart or a mixture of 3D bulk phases restricted to 2 element, layered bulk counterparts. This criteria was used to compute the $\Delta E^{f}_{\text{vac}}$ and $\Delta E^{f}_{\text{ads}}$ energy which determines the stability of the metastable 2D material (Singh, Mathew, et al. 2015; Boland and Singh 2022; Singh and Hennig 2014a; Singh, Zhuang, and Hennig 2014). Appendix D Section D.1 Table D.1 lists the polytype, formation energy, C2DB ID, and spacegroup for each Janus 2D material.

The 3D bulk phase of each 2D material is obtained from the Materials Project (Ong et al. 2013) database, considering only 3D bulk phases that have the (1) lowest $E_{above\ hull}$ and (2) same elemental composition as the 2D material or a mixture of layered, binary bulk phases where a single structure is preferred over a mixture of bulk phases. As noted above, if a 2D material does not have a single structure with a composition matched 3D bulk phase, the energy can be computed from a mixture of 3D bulk phases where the reduced elemental composition of the combined phases are stoichiometrically proportional to the 2D material (Singh, Mathew, et al. 2015). Here, the mixture of 3D bulk phases is limited to 2 layered, binary compounds thereby limiting the computation cost associated with relaxing these structures and the complexity involved in finding the correct stoichiometric mixture of bulk phases. Appendix D Section D.1 Table D.2 lists the Materials Project ID and spacegroup for each 3D bulk phase(s) and discussion for each composition matched 3D bulk phase(s).

In the previous study (Boland and Singh 2022), bulk substrates within the Materials Project (MP) (Ong et al. 2013) database were considered only for structure which were: (a) metallic, (b) cubic phase, (c) single-element, (d) possess a valid ICSD ID (Standards and Technology) (thus been experimentally synthesized), and (e) have the lowest $E_{above\ hull} <$ 0.1 eV/atom. To reduce the computational cost involved

with structure relaxations for the bulk and surface slab structures, only the 19 bulk substrates optimized in the previous study were considered i.e. Al, Au, Ag, Cu, Hf, Ir, Mn, Nd, Ni, Pd, Re, Rh, Sc, Sr, Te, Ti, V, Y, and Zr. All possible combinations of low-index, miller-planes where (h,k,l) can be 0 or 1 were screened for potential lattice matches with all 52 Janus 2D materials. Appendix D Section D.1 Table D.3 lists the Materials Project ID, surface energy, and spacegroup for each substrate.

6.2.2 Heterostructure Screening Criteria

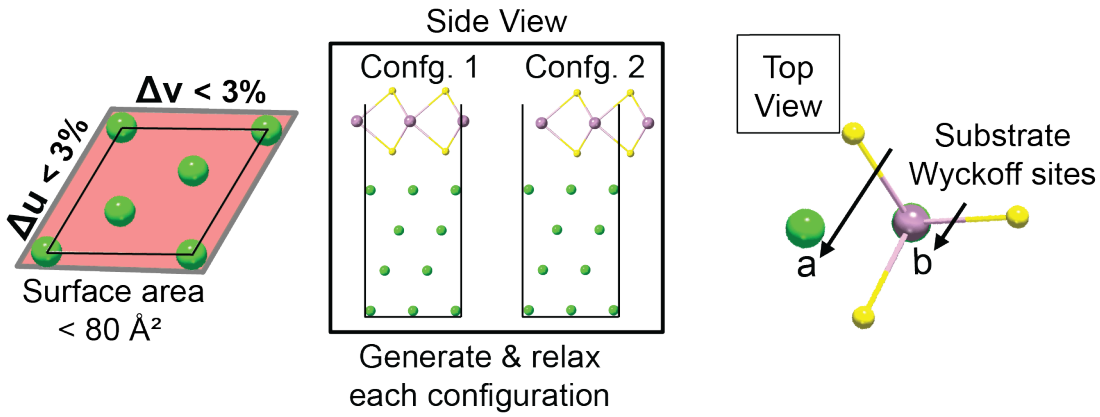

Figure 6.2. Schematic illustrating the criteria for creating heterostructures, the generation of the 2D-substrate heterostructure configurations from the top and side where the top view denotes how the 2D material is aligned on top the substrate surface using the Wyckoff sites.

The maximum heterostructure surface area of $\gamma_{SA} < 80$ Å^2 and applied strain on the 2D material of $\sigma_{2D} < 3\%$ was chosen for the lattice-matching algorithm (Zur and McGill 1984; Mathew et al. 2016). A total of 438 Janus 2D-substrate heterostructure pairs, all on the (111) substrate surface, were identified. 1H-BiClSe was the only 2D material with no resulting heterostructure match for the substrates considered. For each 2D-substrate pair, 2-3 heterostructure configurations are created by enumerating

the high-symmetry points of the 2D material and substrate and aligning the high-symmetry points of the 2D material and substrate z Å apart generating unique heterostructure configures. Of the 438 2D-substrate pairs, 285 pairs (a total of 756 heterostructure configurations) have a negative ΔE^f_{ads} indicating the 2D materials is stabilized by the substrate slab. See Chapter 5 for discussion of ΔE^f_{ads}. Appendix D Section D.1 Table D.4 lists the substrates that produce heterostructures for each Janus 2D material.

6.3 Computational Methodology

6.3.1 DFT

All DFT calculations were performed as outlined in Chapter 5. The density of states (DOS), charge density difference, and Bader analysis (Tang, Sanville, and Henkelman 2009; Henkelman, Arnaldsson, and Jønsson 2006) calculations were automated using the *CMDLElectronicSet* available in the *Hetero2d* package. The k-point mesh grid used to calculate the electronic properties was automated by setting the reciprocal density to 200 with *pymatgen*'s automatic density by volume. The DOS grid was set such that the number of sampling points had an energy spacing of 0.05 eV between each point. The benchmark case, taken with energy spacing of 0.1 eV, 0.05 eV, and 0.01 eV, showed the band gap was converged to within 0.019 eV, see Appendix Figure A.6. The fine grid for the charge density was set to 0.03 Å as this value provided a reasonable compromise between accuracy and computational cost, ∼2.3x the default grid spacing. The benchmark case, taken with grid spacing of 0.1 eV, 0.05 eV, 0.03, and 0.01 eV, showed the Bader charges were converged to within ∼0.004 electrons per

atom, see Appendix Figure D.4. The charge density difference plots demonstrated little change with increased grid spacing, see Appendix Figure D.5. Appendix D Section D.2 discuss the benchmark calculations performed, testing the convergence of the DOS, and the charge density grid spacing used to calculate the Bader charges and charge density difference.

6.3.2 Machine Learning Methodology

The random forest regression models were developed to predict the binding energy (ΔE_{b}), the z-separation distance, and the charge transferred to the 2D material (ΔQ_{2D}) using the *scikit-learn* code (Pedregosa et al. 2011; Buitinck et al. 2013). Random forests are ensemble classifiers which construct multiple independent random tree models that split the input data into subsets on which decision trees are trained to create an aggregate model (Breiman 2001), leading to a more stable and generalizable results. Each model's performance was validated using K-fold cross-validation, and *GridSearchCV* was used to tune the hyper-parameters for each model. Further discussion and figures concerning performance metrics for each model, hyper-parameter tuning, and other feature selection can be found in Appendix D Section D.3.

To generate feature descriptors for the heterostructure and constituent materials, two methods are used: (a) generation of *ab-initio* information using DFT, and (b) using Matminer (Ward et al. 2018) to obtain compositional- and structural-based features from the Materials Project (Ong et al. 2013) database and this study's DFT relaxed 2D-substrate heterostructures, respectively. The data set consists of the lowest energy configuration for each of the 438 Janus 2D-substrate heterostructure pairs optimized using vdW-corrected DFT.

The features added to the data set for machine learning from the 2D-substrate heterostructure DFT calculations were the energetic ($\Delta E^{f}_{\mathrm{vac}}$, ΔE_{b}, substrate's surface energy) and structural (z-separation distance between the 2D material and substrate, δd–see Chapter 5) properties. Additionally, properties from the C2DB were added such as the 2D material's heat of formation and $E_{above\ hull}$. Explicit electronegativity differences between the bottom layer and average of all layers of the 2D material from the substrate's electronegativity were included. Lastly, the charge transferred to the 2D material was added to the list of features.

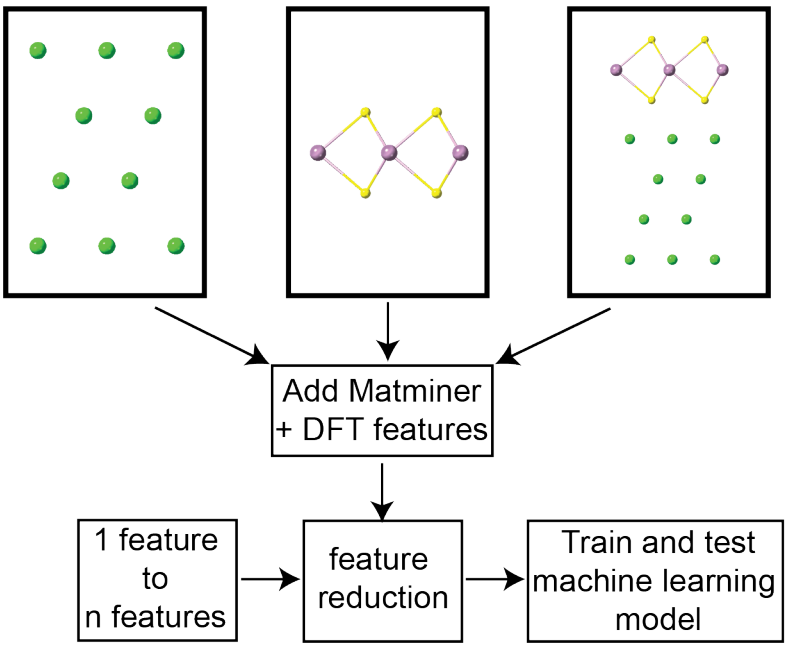

Figure 6.3. Schematic diagram of the steps taken to prepare the data set to perform machine learning for each of the models.

The Matminer (Ward et al. 2018) package was utilized to generate statistical properties (minimum, maximum, range, mean, and standard deviation) for each heterostructure, as well as any additional element based-descriptors attainable from the Materials Project database (Ong et al. 2013). Matminer has over 70 featurizers–a

method used to generate descriptors of a data set in an automated manner–adapted from scientific publications to incorporate a wide variety of materials descriptors.

The *ElementalProperty* featurizer was used to encode fundamental elemental-based properties for each atomic species in the heterostructure and constituent (2D and substrate) structures; this study's data set consists of: the atomic mass, atomic radius, atomic number, group, row, Mendeleev number, bulk modulus, molar volume, thermal conductivity, and electronegativity. The *GlobalSymmetryFeatures* and *MaximumPackingEfficiency* were used to encode structural (packing efficiency of elemental components) and crystallographic (spacegroup number and crystal system) properties. The *ValenceOrbital* featurizer was used to encode electronic properties (weighted fraction of s, p, d, and f valence electrons in each orbital) for the heterostructure.

During featurization the incorporation of the numerous statistical properties can result in many highly correlated and low-variance features. Appendix D Section D.3 provides a detailed description of the steps taken to remove strongly correlated/low-variance features and prevent over-fitting of the models. It should be noted, irregardless of the feature selection method, the overall performance of the model was not significantly impacted. However, the speed of training the model was increased as well as the magnitude of the important features in the list.

6.4 Visualization of the Data Set and Trends

The 2DMSD can guide experimental synthesis identifying the impact the substrate will have on the properties of the 2D material enabling one to select substrates for a wide range of applications from contact materials to a support material within a device. The versatility of the 2DMSD is demonstrated by providing a detailed analysis

of SbTeI demonstrating how the 2DMSD can be used to identify substrates which stabilize the metastable 2D material and characterize the nature of the 2D-substrate interaction.

6.4.1 Energetic Stability of Janus 2D Materials

As discussed in Chapter 5, the first criteria to assess the suitability of a substrate in stabilizing a 2D material is the adsorption formation energy, $\Delta E^f_{\text{ads}} = \Delta E^f_{\text{vac}} - \Delta E_{\text{b}}$ (Singh, Mathew, et al. 2015). The formation energy is given by $\Delta E^f_{\text{vac}} = E_{\text{2D}}/N_{\text{2D}} - E_{\text{3D}}/N_{\text{3D}}$, where E_{2D} is the energy of a 2D material in vacuum, E_{3D} is the energy of the bulk counterpart of the 2D material, and N_{2D} and N_{3D} are the number of atoms in the unit cell of 2D and bulk counterpart, respectively. The binding energy of the 2D-substrate pair is given by $\Delta E_{\text{b}} = (E_{\text{2D}} + E_{\text{S}} - E_{\text{2D+S}})/N_{\text{2D}}$, where $E_{\text{2D+S}}$ is the energy of the 2D material adsorbed on the surface of a substrate, E_{S} is the energy of the substrate slab, E_{2D} is the energy of the free-standing 2D material, and N_{2D} is the number of atoms in the unit cell of the 2D material.

The adsorption formation energy provides a quantitative measure of the stability of the 2D material on the substrate surface. When the difference between ΔE^f_{vac} and ΔE_{b} is less than zero, the substrate stabilizes the 2D material. Figure 6.4 shows the ΔE^f_{ads} for the $1T$ and $1H$ phases of SbTeI. The degree to which the adsorption formation energy of the 2D material is stabilized varies depending on the substrate. Previous reports (Singh and Hennig 2014b; Singh, Hennig, et al. 2015) have indicated that larger ΔE^f_{ads} are more consistent with ionic/covalent interactions, while values closer to zero are more consistent with vdW type interactions. In order to obtain a more quantitative descriptors of the type of bonding between atoms would require

the use of electron localization function (ELF) and analyze the location of the charge density form these calculations. Programs such as Critic2 have been developed to analyze the type of bonding from both Bader and ELF calculations.

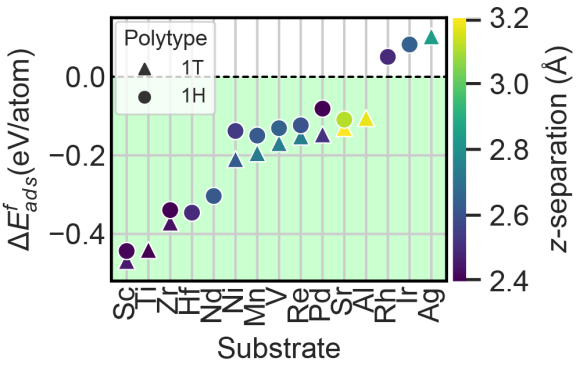

Figure 6.4. Adsorption formation energy (ΔE^f_{ads}) of 1T- (triangle markers) and 1H-SbTeI (circle markers) on the lattice- and symmetry-matched substrate surfaces. The substrates are along the x-axis and the ΔE^f_{ads} for the lowest energy heterostructure configuration is shown on the y-axis. The green shaded region indicates when the substrate stabilizes the 2D material. The color map corresponds to the z-separation distance between the 2D materials and the substrate surface.

Figure 6.4 clearly demonstrates the correlation between the ΔE^f_{ads} and the z-separation distance–which is color mapped on ΔE^f_{ads} values. A majority of the heterostructure pairs with a large adsorption energy have smaller separation distances and the z-separation distances increase as the ΔE^f_{ads} increases towards 0. This dependence is not only true for this example case, but is representative of all 2D-substrate pairs computed in this study and is present even when considering the different configurations for a 2D-substrate pair. Logically, the ΔE_{b} is the actual quantity in the ΔE^f_{ads} which varies with the z-separation distance since the ΔE^f_{ads} is simply the difference between ΔE^f_{vac} and ΔE_{b} which, for a given 2D material, ΔE^f_{vac} is constant. To demonstrate this trend, the ΔE_{b} is plotted in a heatmap sorted from

least (top/left) to greatest (bottom/right) electronegativity of the 2D materials and substrate in Figure 6.5 and Appendix Figure D.1.

For all substrates excluding Rh, Ir, and Ag, the (111) substrate surface stabilizes the SbTeI 2D material. The strength of the interaction varies from relatively weak (near zero) to large values of ∼0.4 eV/atom. As discussed in Section 1.4, vdW interactions are orders of magnitude smaller than ionic/covalent bonds, typically on the order of 0.03 eV/atom and 0.2 eV/atom (Deng et al. 2017). The relatively weak interaction energy correlated with the larger z-separation distances for Sr and Al are encouraging substrates that may facilitate the growth of this Janus 2D material. There is also a clear correlation between the adsorption energy and the z-separation distance.

Since the variations observed in the ΔE_{b} for a single 2D phase over a range of substrates arises from changes in the interactions strength between the 2D material and substrate it may be more useful to analyze ΔE_{b} rather than the adsorption energy to discern trends within the data. Figure 6.5 shows how the ΔE_{b} varies for each 2D material over all the substrate surfaces. One trend of particular interest that is not as apparent from the adsorption energy is the small variation in the ΔE_{b} across the different 2D materials on the same substrate.

In device applications the clear correlation between ΔE_{b} and the z-separation suggests that, in principle, it may be possible to tune the properties of the 2D material by simply changing the support. As the support is changed, the z-separation can be modulated changing the bonding of the 2D material and substrate. To utilize this trend, a clear set of guidelines needs to be established regarding the underlying materials dependent properties that give rise to the variations in interaction strength for any given 2D material and substrate and the impact on the 2D materials properties.

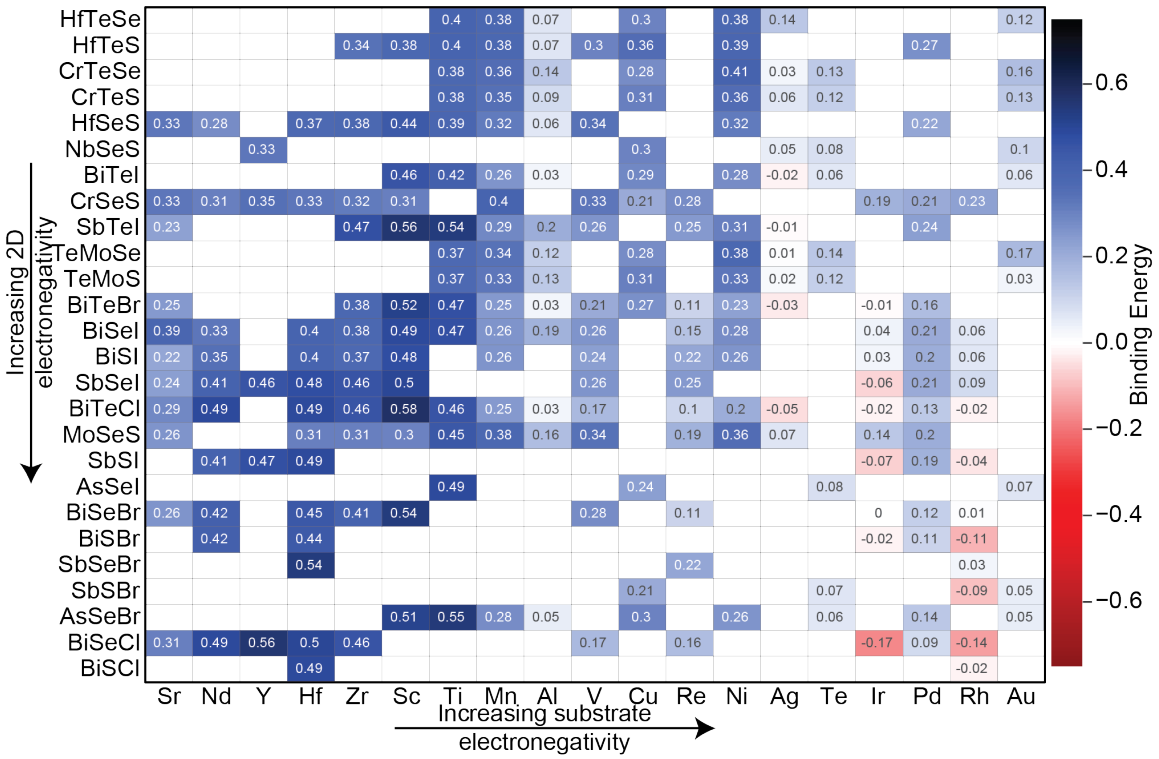

Figure 6.5. Heatmap for the 1T phase of the Janus 2D materials sorted from least (top/left) to greatest (bottom/right) electronegativity of the 2D materials and substrate. The property to sort the 2D and substrate materials were selected using the important features found in the machine learning model. The colormap represents the (ΔE_{b}) binding energy.

6.4.2 Distribution of 2D-Substrate Z-Separation Distances

It is well-known that vdW interactions occur over larger distances than ionic/covalent interactions; as such it serves as a useful screening parameter to search for specific interaction types in large data sets like the 2DMSD. Additionally, it can be useful to compare changes in the z-separation distance for different polytypes with the same composition to understand how the polytype influences the 2D materials properties.

Figure 6.6 shows the z-separation for all 1T polytypes for the 2D-substrate het-

erostructure matches. The x and y axis are sorted from least (top/left) to greatest (bottom/right) electronegativity of the 2D materials and substrate's surface energy. This ordering clearly illustrates that the z-separation is influenced primarily by the substrate's surface energy with little variations for the different 2D materials.

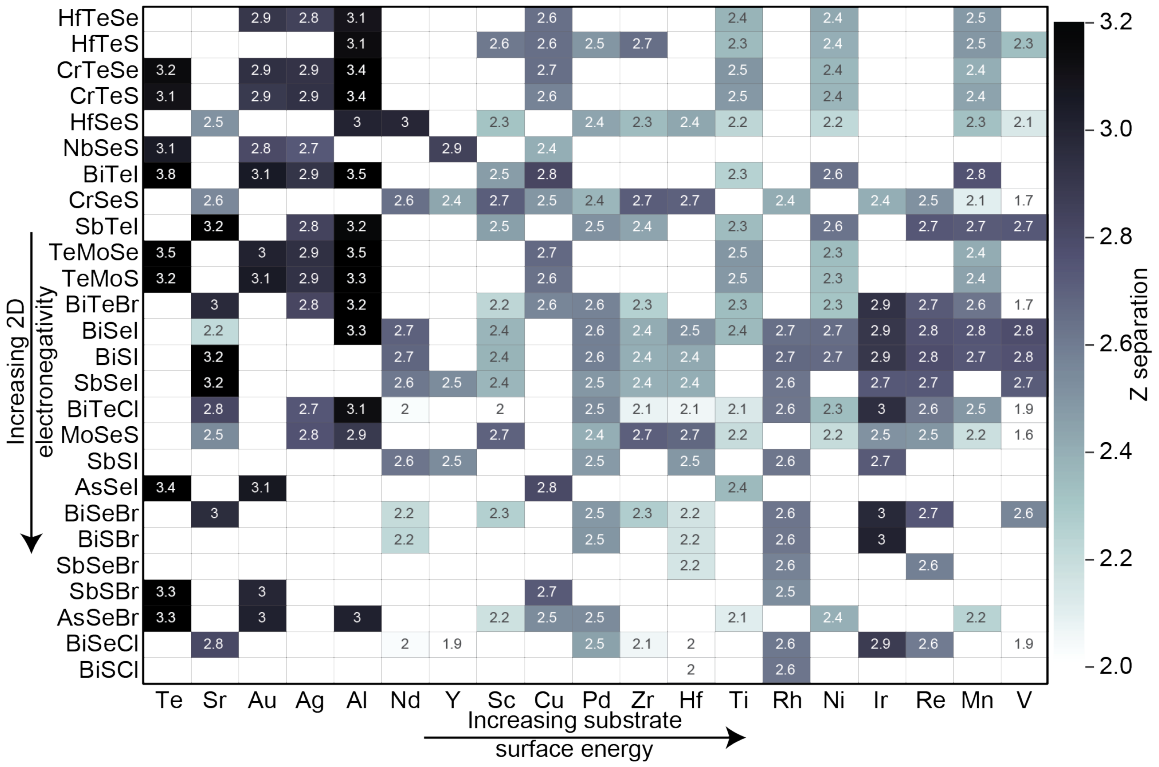

Figure 6.6. Heatmap for the 1T phase of the Janus 2D materials sorted from least (top/left) to greatest (bottom/right) electronegativity of the 2D materials and substrate's surface energy. The colormap represents the z-separation.

Figure 6.7 illustrates the distribution of z-separation distances on various substrates when both polytypes of SbTeI have been calculated. The z-separation distances of the 1H-SbTeI compared with the 1T-SbTeI have slight variations where the 1H polytype has smaller separation distances compared to the 1T polytype.

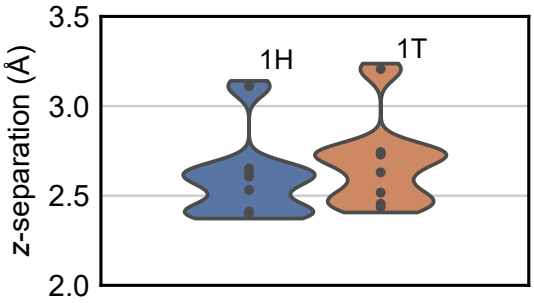

Figure 6.7. The z-separation distance between the top layer of the substrate and the bottom layer of the SbTeI Janus 2D material. The $1T$ polytype displays greater variation in z-separation distance compared to the $1H$ polytype.

6.4.3 Identifying Substrate Induced Charge Transfer

To assess the bonding character between the 2D material and substrate surface slab, the charge transfer is estimated using Bader formalism. The total charge transferred to the 2D material (or substrate) for a given heterostructure (ΔQ_{2D}) is computed using Equation 6.1

$$\Delta Q_i = \sum_{n=1}^{N_i} (Q_n^{combined} - Q_n^{isolated_i}) \tag{6.1}$$

where i represents the sites for the isolated 2D material and substrate, n corresponds to the site index of the n^{th} atom of the 2D-substrate (combined) and two isolated structures, N_i is the total number of sites in i, and $Q_n^{combined/iso}$ is the calculated Bader charge transfer to the n^{th} atom for the combined/isolated system.

The charge transfer on each site is automated using *pymatgen*'s *bader_caller* given by the difference between the total number of electrons on each site and the number of electrons for the neutral atom (ZVAL in the POTCAR). Positive values indicate electron accumulation and negative values indicate electron depletion with respect to the neutral atom. Similarly, positive values of ΔQ_i indicate electron density accumulation to i relative to the isolated structure and visa versa. It is important to

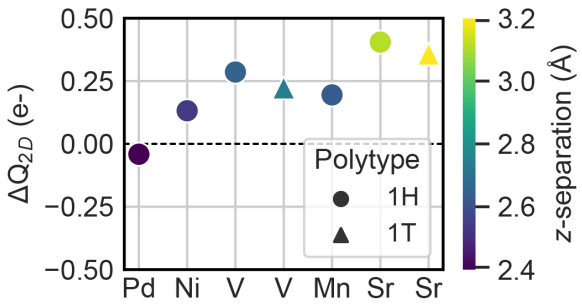

Figure 6.8. The change in the charge transfer to SbTeI (ΔQ_{2D}) is calculated with Bader analysis according to Equation 6.1. Negative (positive) values indicate electron depletion (accumulation) from the 2D material.

note the calculation of Bader volumes is not exact and the total charge on each atom can vary by as much as 1 atomic unit (Choudhuri and Truhlar 2020).

Figure 6.8 shows the charge transferred to the adsorbed 2D SbTeI on the y-axis when placed on the substrate indicated on the x-axis. The marker shapes indicate the polytype of the 2D material - H phases indicated with a $\triangle$ and T phases indicated with a $\circ$. The color map indicates the z-separation distance between the bottom layer of the 2D material and the top layer of the substrate for the given charge transfer. The substrate surfaces are listed from the smallest to the largest electronegativity difference, e.g. the difference between the average electronegativity of the 2D material from that of the substrate. Within each group, the 2D phases are sorted from smallest to largest z-separation distance–for a given 2D-substrate pair each polytype produces the same electronegativity difference. It is particularly interesting to note that for SbTeI, the degree of charge transfer is strongly correlated with the difference between the 2D and substrate electronegativity. However, for all Janus 2D materials, this trend is not always observed. One surprising result is that despite the increasing z-separation distance between the 2D material and substrate surface, there is an increase in the charge transfer.

6.4.4 Electronic Structure and Charge Redistribution of Adsorbed SbTeI

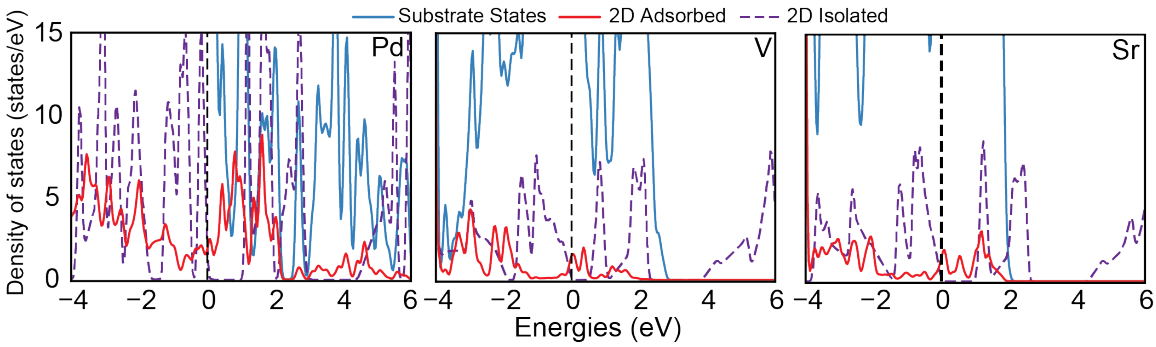

Figure 6.9. Substrate/2D projected density of states (DOS) for the adsorbed (solid red lines) and isolated adsorbed (dashed purple lines) 1H-SbTeI Janus 2D material on three substrate surfaces (solid blue lines) which were selected from the charge transfer graph to have the minimum, median, and maximum charge transfer. The vertical dashed line at x=0 indicates the Fermi level. For each 2D material, the interaction with the substrate surface causes a semiconductor-to-metal transition.

To further understand substrate effects on the 2D material, analysis of the density of states (DOS) can provide insight into the nature of the bonding, and thus the degree of interaction between the 2D material and substrate. The DOS of material is influenced by various properties from electric fields that can arise from the surface, the interaction of dipole moments (such as those present in Janus 2D materials), and/or hybridization or bonding between the 2D material and substrate. These interactions can have a significant impact of the properties of the 2D materials, such as changes in the electronic structure or the charge density distribution.

Figure 6.9 shows the projected DOS for the 1H-SbTeI Janus 2D material on three substrate surfaces. The substrates were selected from the charge transfer figure to represent the minimum, median, and maximum charge transfer seen in Figure 6.8 to assess how charge transfer impacts the electronic DOS in these heterostructures. In all cases, the interaction between the substrate surfaces and 2D material causes

significant changes in the DOS profile of the adsorbed (solid red lines) compared with the isolated-adsorbed 2D DOS profile (dashed purple lines). In general, the DOS profile for the adsorbed Janus 2D material undergoes a semiconductor-to-metal transition and/or hybridizes with the metallic states of the substrate. Even for cases such as the SbTeI on the Sr-(111) substrate surface where the z-separation distance is large ($\sim$3.1 Å), the DOS profile changes from that of the isolated 2D material and a small amount of charge is transferred to the 2D material as indicated by Bader analysis.

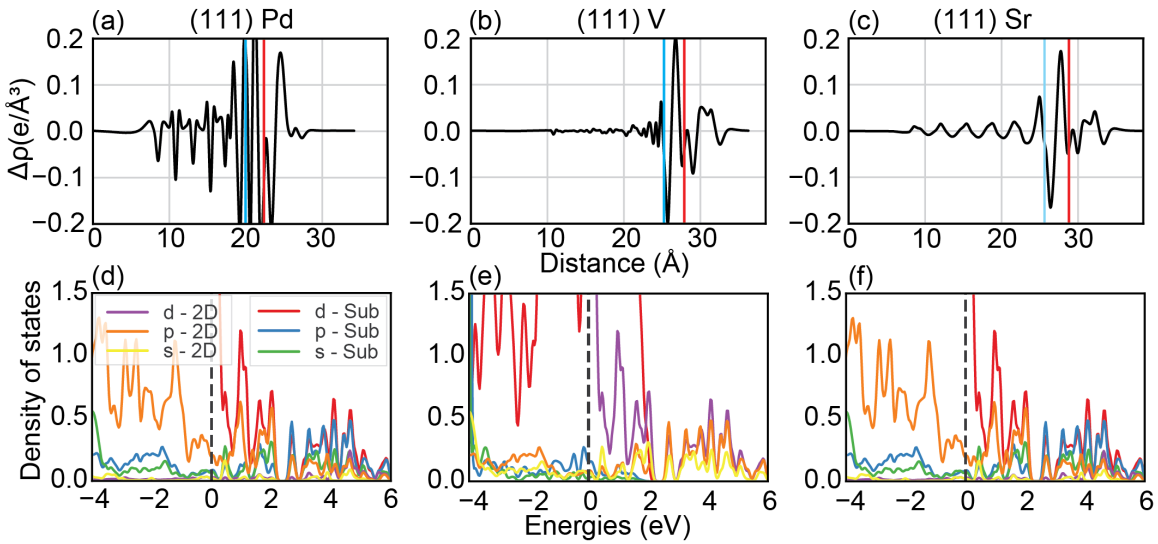

Figure 6.10. (a-c) The z-projected charge density difference for select 2D-substrate heterostructures. The vertical blue line is average z-position of the top-most atomic layer of the substrate surface and the red line is average z-position of the bottom-most atomic layer of the 2D material. The positive values represent electron accumulation and negative values represent electron density depletion with respect to the isolated 2D material and substrate slab. (d-f) The site and orbital projected DOS for the atoms at the interface.

Figure 6.10a-Figure 6.10c show the z-projected charge density difference for the same 1H-SbTeI Janus 2D-substrate pairs discussed above. Appendix Figure D.6 shows the z-projected charge density for the isolated 2D materials and substrate slab and the

combined system which produces the charge density difference. The charge density difference provides a graphical means of determining the charge distribution due to the interactions between the 2D material and the substrate slab. The charge density difference requires three calculations of charge density: (1) the combined system, (2) the isolated 2D material, and (3) the isolated substrate slab in which the atomic position of each atom are frozen in each case. The vertical blue line represent the average z-position of the top-most atomic layer of the substrate surface and the red line is the average z-position of the bottom-most atomic layer of the 2D material. The charge density difference shows significant redistribution of the charge density due to the interactions between the 2D material and substrate surface. Figure 6.10d-Figure 6.10f shows the site and orbital projected DOS for the atoms at the interface. The orbital and site DOS for the combined system demonstrate hybridization as seen by the overlap in the DOS for the 2D and substrate.

6.5 Machine Learning Insights into the Fundamental Factors Governing Janus 2D Heterostructure Properties

As discussed above, the $\Delta E^{f}_{\text{ads}}$, and therefore the ΔE_{b}, varies when the substrate is changed. However, a clear materials properly dependence is not directly discernible from the data plotted in Figure 6.4. Machine learning has been a powerful tool within materials science to identify relationships between materials properties and a relevant (target) quantity of interest such as ΔE_{b}, the z-separation distance, or the charge transferred to the 2D material (ΔQ_{2D}). The machine learning models utilize random forest regression models implemented in the *scikit-learn* code (Pedregosa et al. 2011; Buitinck et al. 2013) to predict the binding energy (ΔE_{b}), z-separation

distance, and charge transfer to the 2D material in 2D-substrate heterostructures. The performance of the three machine learning models and the materials insights elicited for the ΔE_{b}, the z-separation distance, and ΔQ_{2D} for Janus 2D-substrate heterostructures is discussed in the following sections.

6.5.1 Predictive Model for Determining the Binding Energy of Janus 2D Materials

The random forest regression model for predicting the ΔE_{b} has optimal hyperparameters with a K-fold cross-validation value of K=8 used for training and testing. There are a total of 16 features in the feature space and 150 decision trees were used to create the aggregate model. These parameters were found to provide reasonable accuracy and computational speed for training and testing. The random forest regression model has an R^2=0.941, a mean absolute error (MAE) of 0.031 eV/atom, and root mean squared error of 0.041 eV/atom. Figure 6.11a shows the DFT computed and machine learning predicted ΔE_{b} for the data set, where the shaded gray area represents the RMSE from the cross-validation. The machine learning predictions follow the $y = x$ dashed line well, indicating the model provides reasonable accuracy predicting the ΔE_{b} of the 2D-substrate heterostructure pairs which is confirmed by both the MAE and RMSE. The model can be used to tailor the adsorption formation energy for 2D materials by identifying substrates that have a specific range of binding energies for stabilizing the 2D material.

Additionally, from the random forest regression model, important features can be extracted to elucidate the underlying materials properties that were the most predictive features in determining the target property, providing the relative importance of the predictors in the feature space. Figure 6.11b illustrates the top five most predictive

features and their relative importance as determined by the random forest regression model. The most significant features are the electronegativity of the substrate, followed by the z-separation distance, the maximum packing efficiency Ward et al. 2017, atomic number of the substrate, and the average d orbital valence electrons. The maximum packing efficiency of the heterostructure is computed by finding the largest sphere that fits inside a Voronoi cell or, simply, a close-packing sphere model.

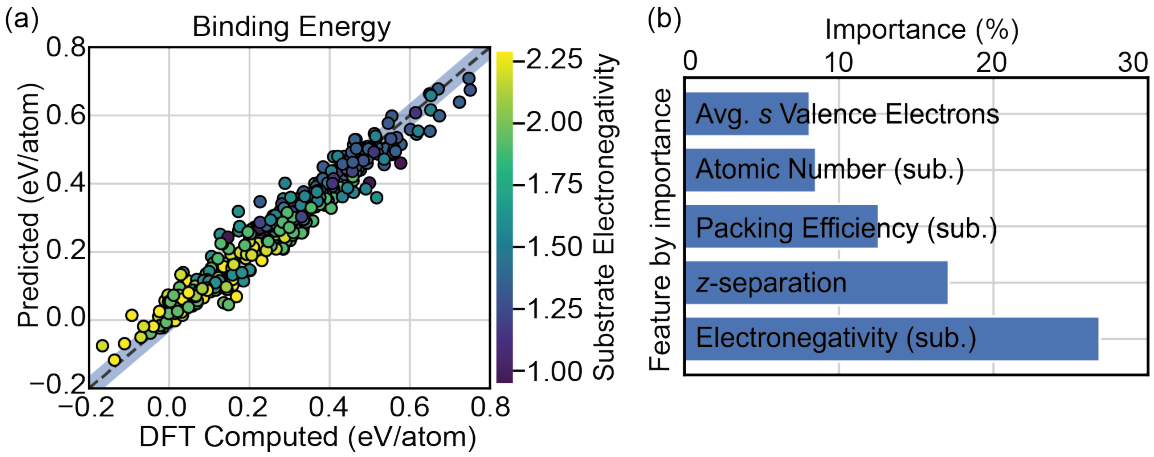

Figure 6.11. (a) Scatter plot compares the DFT computed and machine learning predicted binding energy for the Janus 2D-substrate heterostructure configurations. The color map represents the value of the substrate's electronegativity. The black dashed line indicates the $y = x$ function. The closer the data-points are to this line the better agreement between the predicted binding energy value to the DFT computed value. The shaded region is the RMSE provided by the model. (b) The top five most predictive features ranked by the random forest regression model used to predict the binding energy for each Janus 2D material and substrate surface pair.

The color map in Figure 6.11a represents the substrate's electronegativity which was determined to be the most predictive feature by the model. The binding energy between the 2D material and substrate surface increases as the electronegativity of the substrate decreases. This trend may at first appear counter-intuitive; in general, the more electronegative a material the stronger it will interact with other materials. However, the Janus 2D materials have no dangling bonds to interact with the substrate

surface which many explain despite the relatively large electronegativity variations (∼2-2.6) the primary factor in determining the ΔE_{b} is the substrate's electronegativity. Note, the substrate's radii is a complementary feature for this data set as it produces the same results when used in place of the substrate's electronegativity. The substrate's radius was removed from the feature space, as it was highly correlated and removing the feature did not affect the algorithms performance. However, the high correlation between these two features and equal predictive capability for the machine learning model is an interesting result.

The next important feature in determining the binding energy is the z-separation distance. The smaller the separation between two elements, the more the orbitals interact with each other, which generally increases the stability of the system and results in a strong binding between the 2D material and the substrate. The last three predictive features in Figure 6.11b are the average number of s valence electrons in the heterostructure followed by the substrate's packing efficiency and atomic number. Another similarly predictive feature that was removed from the feature space due to its strong correlation with the s valence electrons is the d valence electrons. The last final two features represent structural factors which determine the density of sites at the surface.

6.5.2 2D-Substrate Z-Separation Distance

The random forest regression model for predicting the z-separation distance has optimal hyper-parameters with a K-fold cross-validation value of K=5 used for training and testing. There are a total of 11 features in the feature space, and 500

decision trees were used to create the aggregate model. These parameters were found to provide reasonable accuracy and computational speed for training and testing.

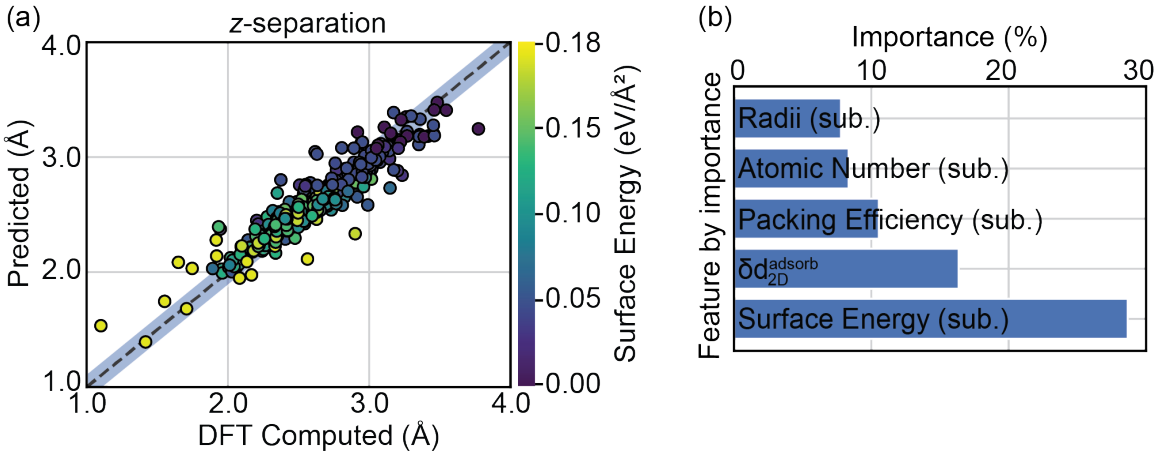

Figure 6.12. (a) Scatter plot comparing the DFT computed and machine learning predicted z-separation distance. The color map represents the value of the substrate's surface energy. The black dashed line indicates the $y = x$ function. The shaded region is the RMSE provided by the model. (b) The bar chart indicates the top five most predictive features return by the random forest regression model for predicting the z-separation distance between the Janus 2D materials and the substrate surface.

The random forest regression model has an R^2=0.876 and a mean absolute error (MAE) of 0.080 Å and root mean squared error of 0.120 Å. Figure 6.12a shows the DFT computed and machine learning predicted z-separation distance for the data set where the shaded gray area represents the RMSE from cross-validation. The machine learning predictions follow the $y = x$ dashed line well for z-separation distance between, indicating the model provides reasonable accuracy predicting the 2D-substrate z-separation distance for heterostructure pairs. This model, combined with the predictive binding energy model, can be used to identify 2D-substrate pairs that have suitable binding energies to successfully stabilize 2D materials and optimal z-separation distances to tune the 2D-substrate interactions.

The data points that deviate significantly from the $y = x$ line are Janus 2D phases

which have disintegrated due to the strong interaction between the bottom layer of particular 2D material-substrate pairs. There are roughly 90 2D-substrate pairs in the data set which have disintegrated or significantly distorted 2D materials. Typically, the bottom layer of the 2D material is composed of Br, Cl, or I and the substrate surface is Hf, Mn, Nd, Sc, Sr, Ti, V, Y, and Zr. If these data point are removed from the data set, the change in the adsorbed 2D materials thickness is no longer needed as a feature in the machine learning model to obtain reasonable accuracy in predicting the z-separation distance.

The top five most-predictive features for the z-separation distance are listed in Figure 6.12b as determined by the random forest regression model. The most significant features are the substrate's surface energy, followed by the change in the adsorbed 2D materials thickness, the maximum packing efficiency Ward et al. 2017, atomic number, and radii of the substrate.

The color map overlaid on each data-point in Figure 6.12a represents the value of the substrate's surface energy. In general, as the surface energy increases, the z-separation distance between the 2D material and substrate surface decreases. This may be due to the larger number of dangling bonds present at the surface increasing the likelihood of the 2D materials and substrate interacting.

The next important feature in determining the z-separation distance is the thickness changes in the adsorbed 2D material. This features is likely important in determining the target property since larger values indicate the 2D material has disintegrated on the substrate surface due to the strong interactions between the bottom layer of the 2D material and top layer of the substrate surface. It is worth noting that removing the disintegrated features does not significantly change the important features, only the order of importance for the last 3 features.

6.5.3 Charge Transfer to the 2D Material

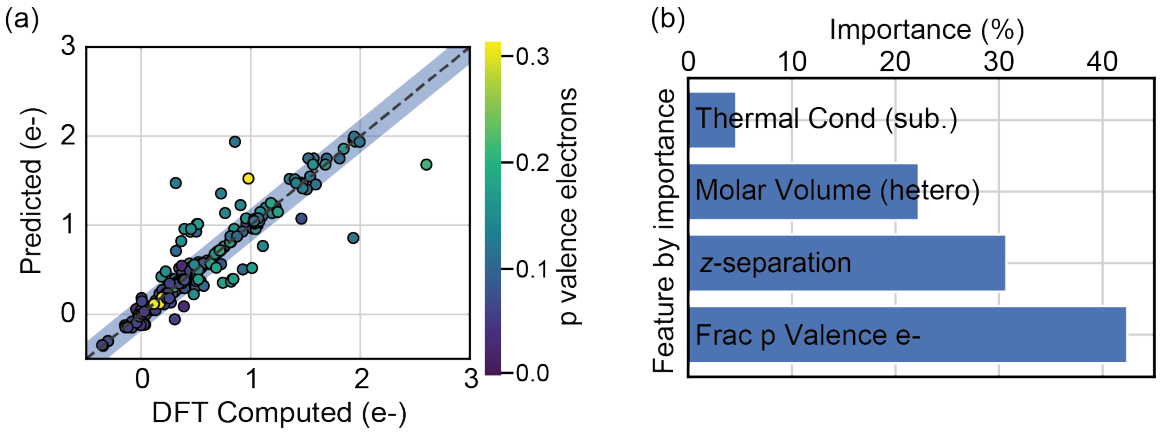

Figure 6.13. (a) Scatter plot comparing the DFT computed and machine learning predicted ΔQ_{2D}. The color map represents the fractional p valence electrons within the heterostructure. The black dashed line indicates the $y = x$ function. The shaded region is the RMSE provided by the model. (b) The bar chart indicates predictive capability of each feature in the feature space.

The random forest regression model for predicting the ΔQ_{2D} has optimal hyper-parameters with a K-fold cross-validation value of K=3 used for training and testing. There are a total of 4 features in the feature space, and 50 decision trees were used to create the aggregate model. These parameters were found to provide reasonable accuracy and computational speed for training and testing.

The random forest regression model has an R^2=0.845 and a mean absolute error (MAE) of 0.098 electrons and root mean squared error of 0.203 electrons. Figure 6.13a shows the DFT computed and machine learning predicted ΔQ_{2D} for the data set where the shaded gray area represents the RMSE from cross-validation. The machine learning predictions follow the $y = x$ dashed line well for ΔQ_{2D} reasonable accuracy in predicting the ΔQ_{2D} for the heterostructures.

The most-predictive features for the ΔQ_{2D} are listed in Figure 6.13b as determined

by the random forest regression model. The color map overlaid on each data-point in Figure 6.13a represents the value of the fraction of p valence electrons which is the most significant features is the fraction of p valence electrons.

6.6 Summary

In summary, a new database 2DMSD for 2D-substrate heterostructures and their associated properties is presented. Analyzing the database, a correlation between the large adsorption energies and smaller separation distances was found and as the adsorption energy approaches zero the z-separation distance increases. One surprising result is that despite the increasing z-separation distance between the 2D material and substrate surface, there is an increase in the charge transfer. The interaction between the metal substrate and the Janus 2D materials results in metal-to-semiconductor transition and wide-spread changes in the electronic structure.

The wide range of potential chemical (bonding, elemental, and structural) interactions among the 2D materials and substrates make this problem a perfect candidate system to utilize machine learning to identify the key materials properties which determine the interfacial interaction strength which is discussed in the next section. These calculations provide some insight into the underlying mechanism that results in variations in the ΔE_{b} z-separation, and charge transfer when changing the substrate or 2D material.

The regression models indicate that the most critical property in determining the binding energy, z-separation distance, and charge transferred to the 2D material are the substrate's electronegativity, surface energy, and fraction of p valence electrons, respectively. The binding energy between the 2D material and substrate surface

increases as the electronegativity of the substrate decreases. Using the machine learning model for predicting the z-separation distance, it is observed that as the surface energy increases the z-separation distance between the 2D material and substrate surface decreases. This may be due to the larger number of dangling bonds present at the surface increasing the likelihood of the 2D materials and substrate interacting. The ΔQ_{2D} regression model indicated that the fraction of p valence electrons is the most-predictive feature in determining the charge transferred between the 2D material and substrate.

Chapter 7

FUTURE WORK

7.1 Conclusions and Outlook

Unsurprisingly, the nanoscale interfacial regions and their properties are inherently intertwined and modulated by the properties of the adjacent materials. The presence of defects (point, planar, or surface defects) can result in large disparities in the observed interfacial properties even within a single sample. The objective of this research was to apply DFT, high-throughput methodologies, and machine learning to model interfacial systems to predict their properties and discover interface-property relationships providing guidance towards future studies and the optimization of these regions.

7.1.1 Conclusions and Outlook for Grain Boundary Studies

In polycrystalline doped ceria, the randomly oriented grain boundaries (William J Bowman et al. 2017) produce large variations in the grain boundary character and markedly different interfacial properties. Even macroscopic considerations such as grain size effects (Rupp, Infortuna, and Gauckler 2006; Lee, Prinz, and Cai 2013; Hwang and Mason 1998) and space charge layers (Maier 1995; De Souza, Ramadan, and Hörner 2012; Shirpour et al. 2012; S. Kim et al. 2016) play an important role in altering the interfacial properties in these materials. As discussed in Chapter 1 Section 1.3, recent experimental and computational progress has been made towards

understanding the interfacial properties of grain boundaries, however, the effects of high local solute compositions was previously unexplored. This dissertation work utilized DFT to provide a means to systematically study grain boundary properties subject to local composition changes and assess how these variations modulate interfacial properties.

Using DFT with the GGA+U functional, the $\Sigma 3$ (111)/[$\bar{1}$01] and $\Sigma 3$ (121)/[$\bar{1}$01] grain boundary structures were systematically doped with AEM solutes to assess the role high solute concentrations play in modulating the interfacial properties of ceria. These AEM were selected to capture solute size and compositional trends in the properties of two high-angle grain boundaries. The energetic stability of the $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary was found to have the lowest grain boundary energy whereas doping with AEM solutes produced low grain boundary energies when the ionic radii is similar or larger than the host solute Ce i.e. Ca, Sr, and Ba. Additionally, the energetic stability associated with accommodating the solutes depend on the local atomic grain boundary structure and the coordination number of the substitutional site. The less coherent $\Sigma 3$ (121)/[$\bar{1}$01] grain boundary structures demonstrated less site and solute variation in its interfacial properties compared with the coherent $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary.

To explore the relationship between grain boundary solute concentration and oxygen ionic conductivity reported by William J Bowman et al. (2017), the near grain boundary region of the Ca-doped $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary was systematically investigated assessing the segregation energy and migration energy of oxygen vacancies. In this work the presence of Ca ions were found to not significantly alter the segregation behavior of oxygen vacancies compared with the undoped system; the difference in segregation energy at each oxygen plane is at most 0.4 eV and 0.6 eV for the first and

second Ca-doped system, respectively. However, the segregation behavior of oxygen vacancies within the Ca-doped grain boundary core has a smoother potential energy landscape for the first Ca-doped system compared to the undoped system while the second Ca-doped system exhibits a stronger trapping of oxygen vacancies within the grain boundary core.

The presence of the Ca solute ion does have a significant impact on the migration energy of the oxygen vacancy in the grain boundary. The presence of the Ca solute results in a significant reduction in the oxygen vacancy migration energy within and towards the grain boundary core when compared to migration energy values found in the bulk and in literature for this particular grain boundary. However, depending on the site of the Ca ion within the grain boundary core the migration energy could be significantly lower compared to the undoped system. Despite this, these results provide critical insight into the role that high local solute concentrations play in modulating the migration energy of oxygen vacancies improving the understanding of cross boundary oxygen transport. Overall, the segregation of oxygen vacancies towards the grain boundary core is favorable both thermodynamically and kinetically. However, the presence of larger migration energy barriers for certain oxygen planes coupled with an appreciate increase in the segregation energy would inhibit the cross boundary migration decreasing the local ionic conductivity in certain regions.

In future studies, it would be interesting to assess the change in the potential energy landscape and migration energy with additional solutes and oxygen vacancies present in the bulk and the grain boundary over a wider range of misorientation angles. This would provide further insights into how the migration energy landscape would change with increased oxygen vacancy concentrations and the impact this would have on the oxygen vacancy stability within the grain boundary region. However, the increased

system size required to model the nominal solute/oxygen vacancy concentrations in the bulk precludes the exploration of such phenomena using plane-wave based DFT methods. The issue is further complicated by the technical difficulties one would certainly experience with the increased oxygen vacancy concentrations and the necessity to identity Ce sites which the electrons would localize on. Additionally, more elaborate solute profiles at the grain boundary would be of interest to further explore the relationship between migration energy, misorientation, composition, and concentration in other grain boundaries. These future studies would be more well suited for molecular dynamics simulations. However, considering the considerable amount of time and resources required to generate new interatomic potentials and the limited applicability of current interatomic potentials developed to describe bulk systems, this is a challenging avenue to pursue. More recently, machine learning interatomic potentials have gained much attention for their improved accuracy, smaller training set requirements, and flexibility, making them a potential tools to computationally explore more realistic system sizes and composition profiles of grain boundaries.

7.1.2 Computational Synthesis of 2D-Substrate Heterostructured Materials

2D materials show promise for a wide range of potential applications from biological to energy generation technologies. Considering the vast chemical and elemental diversity of the >5,000 predicted 2D materials, it is unsurprising that realization of these theoretically predicted materials has been slow. The interplay between the 2D material, substrate surface, and presence of defects (in the 2D material or the substrate) can dramatically alter the unique properties of the 2D material (Y.-H. Lee et al. 2013; Tang et al. 2021; Singh and Hennig 2014b); this can pose significant

challenges in the successful synthesis and realization of the many as-yet-hypothesized 2D materials when the nature of the 2D-substrate interaction is not known. There are currently no clear-cut guidelines from which to synthesize many of these hypothesized 2D materials, or understand how many of these 2D materials will interact with their environment. The need for a detailed understanding of the substrate-assisted modification of the physical (energetic, structural, chemical, etc.) and electronic (charge density, electronic structure, etc.) properties of post-adsorbed 2D materials. The development and utilization of high-throughput computational workflows can be serve as a means to elucidate a detailed understanding of the interfacial properties in these systems.

In this dissertation project, an open-source workflow package discussed in Chapter 5 was developed to automate the generation of 2D-substrate heterostructures, the creation of DFT input files, the submission and monitoring of computational jobs on supercomputing facilities, and the storage of relevant parameters alongside the post-processed results in a MongoDB database. This workflow dramatically increases the number of calculations possible while decreasing the total number of human hours required to perform the calculations. A benchmark set of four candidate 2D materials and 50 potential substrates was performed to demonstrate the capabilities of the package. The results show a total of 9 (110) and 15 (111) substrate surfaces that stabilize the 2D materials. Analysis of the 2D material's thickness shows most adsorbed 2D materials thickness has increased substantially, which may indicate the interaction between the 2D material and the substrate surface is covalent/ionic which is further supported by the DOS and Bader charge analysis showing a semiconductor-to-metal transition.

To unearth underlying materials properties which govern the stability of 52 Janus

2D materials on 19 metallic substrates, the *Hetero2d* package was utilized generating 438 2D-substrate pairs. Additionally to distribute the data, the 2DMSD database was created and released for 2D-substrate heterostructures and their associated properties. Analyzing the database, a correlation between the large adsorption energies and small z-separation distances was found where adsorption energy approaching zero tended to have larger z-separation distance. Analyzing the charge transfer to the 2D materials reveals that utilizing only the z-separation distance between the 2D material and substrate surface is not sufficient to determine the degree of charge transfer. In all cases, the interaction between the metal substrate and the Janus 2D material results in metal-to-semiconductor transition and wide-spread changes in the electronic structure of the 2D material. These substrates could serve as good electrical contact materials for this class of 2D materials.

Machine learning models were utilized to generate predictive models and elucidate the fundamental materials properties which determine the binding energy and z-separation distance. The regression models indicate that the most critical property in determining the binding energy and z-separation distance are the substrate's electronegativity and surface energy, respectively. These models can be utilized to guide more accurate DFT calculations and aid in the selection of substrates with target properties.

The combination of high-throughput workflows and machine learning methods facilitate a unique opportunity to rapidly generate and discover new material-property relationships at a pace that was not previously possible. The availability of computation workflows opens new means to advance science beyond a single domain. Utilizing additional community workflows it is possible to easily expand and incorporate new types of simulations, transformations, and input data sets. The 2D materials can

be functionalized using ligands or tuned by incorporating point defects in the 2D material. Future studies could be directed towards:

- Exploring the relationship between the substrate surface plane (miller plane, termination, surface reconstruction) and the 2D materials properties and their interfacial interactions.
- Tuning the properties of 2D heterostructures via functionalization using ligands, defects, or add-atoms.
- Fundamental studies which develop guiding principles for a wider range of 2D materials is still needed to understand the full range of chemical interaction between an arbitrary 2D materials and substrate.

REFERENCES

Ahn, Kiyong, Yong-Chae Chung, Kyung Joong Yoon, Ji-Won Son, Byung-Kook Kim, Hae-Weon Lee, and Jong-Ho Lee. 2014. "Lattice-strain effect on oxygen vacancy formation in gadolinium-doped ceria." *Journal of Electroceramics* 32 (1): 72–77. https://doi.org/10.1007/s10832-013-9844-6.

Aidhy, Dilpuneet S. 2016. "Oxygen diffusion in ThO_2 - CeO_2 and ThO_2 - UO_2 solid solutions from atomistic calculations." *Physical Chemistry Chemical Physics* 18 (22): 15019–15024. https://doi.org/10.1039/C6CP01214K.

Aidhy, Dilpuneet S., Y. W. Zhang, and William J. Weber. 2014. "Impact of segregation energetics on oxygen conductivity at ionic grain boundaries." *Journal of Materials Chemistry A* 2 (6): 1704–1709. https://doi.org/10.1039/c3ta14128d.

Aidhy, Dilpuneet S., Yanwen Zhang, and William J. Weber. 2013. "Stabilizing nanocrystalline grains in ceramic-oxides." *Physical Chemistry Chemical Physics* 15:18915. https://doi.org/10.1039/c3cp53052c.

Akiyama, Toru, Takahiro Kawamura, and Tomonori Ito. 2021. "Computational discovery of stable phases of graphene and h-BN van der Waals heterostructures composed of group III-V binary compounds." *Applied Physics Letters* 118 (2). https://doi.org/10.1063/5.0032452.

Amani, Matin, Matthew L. Chin, Alexander L. Mazzoni, Robert A. Burke, Sina Najmaei, Pulickel M. Ajayan, Jun Lou, and Madan Dubey. 2014. "Growth-substrate induced performance degradation in chemically synthesized monolayer MoS_2 field effect transistors." *Applied Physics Letters* 104 (20): 203506. https://doi.org/10.1063/1.4873680.

An, J., J. S. Park, A. L. Koh, H. B. Lee, H. J. Jung, J. Schoonman, R. Sinclair, T. M. Gur, and F. B. Prinz. 2013. "Atomic Scale Verification of Oxide-Ion Vacancy Distribution near a Single Grain Boundary in YSZ." *Scientific Reports* 3 (6). https://doi.org/10.1038/srep02680.

Andersson, D. A., S. I. Simak, N. V. Skorodumova, I. A. Abrikosov, and B. Johansson. 2006. "Optimization of ionic conductivity in doped ceria." *Proceedings of the National Academy of Sciences* 103:3518–3521. https://doi.org/10.1073/pnas.0509537103.

Arora, Gaurav, and Dilpuneet S. Aidhy. 2017. "Segregation and binding energetics at grain boundaries in fluorite oxides." *Journal of Materials Chemistry A* 5:4026–4035. https://doi.org/10.1039/c6ta09895a.

Ashton, Michael, Joshua Paul, Susan B. Sinnott, and Richard G. Hennig. 2017. "Topology-scaling identification of layered solids and stable exfoliated 2D materials." *Physical Review Letters* 118 (10): 106101. https://doi.org/PhysRevLett.118.106101.

Avila-Paredes, H. J., K. Choi, C. T. Chen, and S. Kim. 2009. "Dopant-concentration dependence of grain-boundary conductivity in ceria: A space-charge analysis." *Journal of Materials Chemistry* 19 (27). https://doi.org/10.1039/b904583j.

Avila-Paredes, H. J., and S. Kim. 2006. "The effect of segregated transition metal ions on the grain boundary resistivity of gadolinium doped ceria: Alteration of the space charge potential." *Solid State Ionics* 177 (35-36): 3075–3080. https://doi.org/10.1016/j.ssi.2006.08.017.

Balducci, Gabriele, Jan Kašpar, Paolo Fornasiero, Mauro Graziani, M Saiful Islam, and Julian D Gale. 1997. "Computer Simulation Studies of Bulk Reduction and Oxygen Migration in CeO_2 - ZrO_2 Solid Solutions." *The Journal of Physical Chemistry B* 101 (10): 1750–1753. https://doi.org/10.1021/jp962530g.

Barnard, A. S., B. Motevalli, A. J. Parker, J. M. Fischer, C. A. Feigl, and G. Opletal. 2019. "Nanoinformatics, and the big challenges for the science of small things." *Nanoscale* 11 (41): 19190–19201. https://doi.org/10.1039/C9NR05912A.

Bishop, Sean R., Todd S. Stefanik, and Harry L. Tuller. 2011. "Electrical conductivity and defect equilibria of $Pr_{0.1}Ce_{0.9}O_{2-\delta}$." *Physical Chemistry Chemical Physics* 13 (21): 10165–10173. https://doi.org/10.1039/C0CP02920C.

Blöchl, P. E. 1994. "Projector augmented-wave method." *Physical Review B* 50 (24): 17953–17979. https://doi.org/10.1103/PhysRevB.50.17953.

Blonsky, Michael N., Houlong L. Zhuang, Arunima K. Singh, and Richard G. Hennig. 2015. "Ab initio prediction of piezoelectricity in two-dimensional materials." *ACS Nano* 10 (9): 9885–9891. https://doi.org/10.1021/acsnano.5b03394.

Bokov, Arseniy, Jeffery A. Aguiar, Matthew L. Gong, Alexey Nikonov, and Ricardo H. R. Castro. 2018. "A Strategy to Mitigate Grain Boundary Blocking in Nanocrystalline Zirconia." *Journal of Physical Chemistry C* 122:26344–26352. https://doi.org/10.1021/acs.jpcc.8b08877.

Boland, Tara M., Peter Rez, Peter A. Crozier, and Arunima K. Singh. 2021. "Impact of Aliovalent Alkaline-Earth metal solutes on Ceria Grain Boundaries: A density functional theory study." *Acta Materialia* 205:116481. https://doi.org/10.1016/j.actamat.2020.11.023.

Boland, Tara M., and Arunima K. Singh. 2022. "Computational synthesis of 2D materials: A high-throughput approach to materials design." *Computational Materials Science* 207:111238. https://doi.org/10.1016/j.commatsci.2022.111238.

Bondevik, Tarjei. 2019. "Grain boundaries in a $BaZrO_3$-based proton conductor. A theoretical and experimental study on atomic scale." PhD dissertation, University of Oslo.

Bondevik, Tarjei, Tor S. Bjørheima, and Truls Norbya. 2020. "Assessing common approximations in space charge modeling to estimate the proton resistance across grain boundaries in Y-doped $BaZrO_3$." *Physical Chemistry Chemical Physics* 22. https://doi.org/10.1039/c9cp06625j.

Bondevik, Tarjei, Akihide Kuwabara, and Ole Martin Løvvik. 2019. "Application of machine learning-based selective sampling to determine $BaZrO_3$ grain boundary structures." *Computational Materials Science* 164 (June): 57–65. https://doi.org/10.1016/j.commatsci.2019.03.054.

Born, M., and R. Oppenheimer. 1927. "Zur Quantentheorie der Molekeln." *Annalen der Physik* 389:457–484. https://doi.org/10.1002/andp.19273892002.

Bowler, Jiří Klimeš David R., and Angelos Michaelides. 2011. "Van der Waals density functionals applied to solids." *Physical Review B* 83 (19): 195131. https://doi.org/10.1103/PhysRevB.83.195131.

Bowman, William J, Madeleine N Kelly, Gregory S Rohrer, Cruz A Hernandez, and Peter A Crozier. 2017. "Enhanced ionic conductivity in electroceramics by nanoscale enrichment of grain boundaries with high solute concentration." *Nanoscale* 9:17293–17302. https://doi.org/10.1039/C7NR06941C.

Bowman, William J., Amith Darbal, and Peter A. Crozier. 2020. "Linking macroscopic and nanoscopic ionic conductivity: Towards a new semi-empirical framework for characterizing grain boundary conductivity in polycrystalline ceramics." *ACS Applied Materials & Interfaces* 12:507–517. https://doi.org/10.1021/acsami.9b15933.

Bowman, William J., Jiangtao Zhu, Renu Sharma, and Peter A. Crozier. 2015. "Electrical conductivity and grain boundary composition of Gd-doped and Gd/Pr

co-doped ceria." *Solid State Ionics* 272:9–17. https://doi.org/10.1016/j.ssi.2014.12.006.

Bowman, William John. 2016. *Correlating Nanoscale Grain Boundary Composition with Electrical Conductivity in Ceria.* Arizona State University.

Braithwaite, J. Spencer, and Peter Rez. 2005. "Grain boundary impurities in iron." *Acta Materialia* 53 (9): 2715–2726. https://doi.org/10.1016/j.actamat.2005.02.033.

Breiman, Leo. 2001. "Random Forests." *Machine Learning* 45:5–32. https://doi.org/10.1023/A:1010933404324.

Britannica, Encyclopedia. 2022. *Interface.* https://www.britannica.com/science/interface-physics. [Online; accessed 26. Mar. 2022], March.

Brothers, Edward N, Artur F Izmaylov, Jacques O Normand, Verónica Barone, and Gustavo E Scuseria. 2008. "Accurate solid-state band gaps via screened hybrid electronic structure calculations." *The Journal of Chemical Physics* 129 (1): 011102. https://doi.org/10.1063/1.2955460.

Broyden, Charles G. 1965. "A Class of Methods for Solving Nonlinear Simultaneous Equations." *Mathematics of Computation* 19 (92): 577–593. https://doi.org/10.2307/2003941.

Buitinck, Lars, Gilles Louppe, Mathieu Blondel, Fabian Pedregosa, Andreas Mueller, Olivier Grisel, Vlad Niculae, et al. 2013. "API design for machine learning software: experiences from the scikit-learn project." In *ECML PKDD Workshop: Languages for Data Mining and Machine Learning,* 108–122.

Burke, Kieron. 2012. "Perspective on density functional theory." *The Journal of chemical physics* 136. https://doi.org/10.1063/1.4704546.

Cantwell, Patrick R., Ming Tang, Shen J. Dillon, Jian Luo, Gregory S. Rohrer, and Martin P. Harmer. 2014. "Grain Boundary Complexions." *Acta Materialia* 62:1–48. https://doi.org/10.1016/j.actamat.2013.07.037.

Cargnello, Matteo, Vicky V. T. Doan-Nguyen, Thomas R. Gordon, Rosa E. Diaz, Eric A. Stach, Raymond J. Gorte, Paolo Fornasiero, and Christopher B. Murray. 2013. "Control of metal nanocrystal size reveals metal-support interface role for ceria catalysts." *Science* 341 (6147): 771–773. https://doi.org/10.1126/science.1240148.

Chan, Maria KY, and Gerbrand Ceder. 2010. "Efficient Band Gap Predictions for Solids." *Physical Review Letters* 105 (19). https://doi.org/10.1103/PhysRevLett.105.196403.

Chaves, A, Javad G Azadani, Hussain Alsalman, DR Da Costa, R Frisenda, AJ Chaves, Seung Hyun Song, Young Duck Kim, Daowei He, Jiadong Zhou, et al. 2020. "Bandgap engineering of two-dimensional semiconductor materials." *npj 2D Materials and Applications* 4, no. 29 (August): 1–21. https://doi.org/10.1038/s41699-020-00162-4.

Chen, Wei, Elton J.G. Santos, Wenguang Zhu, Efthimios Kaxiras, and Zhenyu Zhang. 2013. "Tuning the electronic and chemical properties of monolayer MoS_2 adsorbed on transition metal substrates." *Nano Letters* 13 (2): 509–514. https://doi.org/10.1021/nl303909f.

Cheng, Jianli, Jian Luo, and Kesong Yang. 2018. "Aimsgb: An algorithm and open-source python library to generate periodic grain boundary structures." *Computational Materials Science* 155:92–103. https://doi.org/10.1016/j.commatsci.2018.08.029.

Choudhary, Kamal, Kevin F. Garrity, Andrew C. E. Reid, Brian DeCost, Adam J. Biacchi, Angela R. Hight Walker, Zachary Trautt, et al. 2020. "The joint automated repository for various integrated simulations (JARVIS) for data-driven materials design." *npj Computational Materials* 6 (1): 173. https://doi.org/10.1038/s41524-020-00440-1.

Choudhuri, Indrani, and Donald G. Truhlar. 2020. "Calculating and Characterizing the Charge Distributions in Solids." *Journal of Chemical Theory and Computation* 16 (9): 5884–5892. https://doi.org/10.1021/acs.jctc.0c00440.

Ci, Lijie, Li Song, Chuanhong Jin, Deep Jariwala, Dangxin Wu, Yongjie Li, Anchal Srivastava, et al. 2010. "Atomic layers of hybridized boron nitride and graphene domains." *Nature Materials* 9 (5): 430–435. https://doi.org/10.1038/nmat2711.

Ciszewski, Regina. 2012. "Fuel Cells: Optimizing Grain Boundaries in $BaZrO_3$ and Modeling Proton Conduction Using Kinetic Monte Carlo."

Curtarolo, Stefano, Wahyu Setyawan, Gus L.W. Hart, Michal Jahnatek, Roman V. Chepulskii, Richard H. Taylor, Shidong Wang, et al. 2012. "AFLOW: An automatic framework for high-throughput materials discovery." *Computational Materials Science* 58:218–226. https://doi.org/10.1016/j.commatsci.2012.02.005.

Dawson, James A., Hungru Chen, and Isao Tanaka. 2015. "First-Principles calculations of oxygen vacancy formation and metallic behavior at a beta-MnO_2 grain boundary." *ACS Applied Materials and Interfaces* 7 (3): 1726–1734. https://doi.org/10.1021/am507273c.

De Souza, Roger A, Amr Ramadan, and Stefanie Hörner. 2012. "Modifying the barriers for oxygen-vacancy migration in fluorite-structured CeO_2 electrolytes through strain: a computer simulation study." *Energy & Environmental Science* 5 (1): 5445–5453. https://doi.org/10.1039/C2EE02508F.

Deng, Shikai, Enlai Gao, Zhiping Xu, and Vikas Berry. 2017. "Adhesion Energy of MoS_2 Thin Films on Silicon-Based Substrates Determined via the Attributes of a Single MoS_2 Wrinkle." *ACS Publications* 9 (8): 7812–7818. https://doi.org/10.1021/acsami.6b16175.

Dholabhai, Pratik P, Jeffery A Aguiar, Longjia Wu, Terry G Holesinger, Toshihiro Aoki, Ricardo HR Castro, and Blas P Uberuaga. 2015. "Structure and segregation of dopant–defect complexes at grain boundaries in nanocrystalline doped ceria." *Physical Chemistry Chemical Physics* 17:15375–15385. https://doi.org/10.1039/C5CP02200B.

Dholabhai, Pratik P., James B Adams, Peter Crozier, and Renu Sharma. 2011. "In search of enhanced electrolyte materials: a case study of doubly doped ceria." *Journal of Materials Chemistry* 21, no. 47 (November): 18991–18997. https://doi.org/10.1039/C1JM14417K.

Dholabhai, Pratik P., and James B. Adams. 2012. "A blend of first-principles and kinetic lattice Monte Carlo computation to optimize samarium-doped ceria." *Journal of Materials Science* 47:7530–7541. https://doi.org/10.1007/s10853-012-6398-y.

Dholabhai, Pratik P., James B. Adams, Peter Crozier, and Renu Sharma. 2010a. "A density functional study of defect migration in gadolinium doped ceria." *Physical Chemistry Chemical Physics* 12:7904. https://doi.org/10.1039/B924534K.

———. 2010b. "Oxygen vacancy migration in ceria and Pr-doped ceria: A DFT+U study." *Journal of Chemical Physics* 132:1–8. https://doi.org/10.1063/1.3327684.

Dholabhai, Pratik P., Shahriar Anwar, James B Adams, Peter Crozier, and Renu Sharma. 2011. "Kinetic lattice Monte Carlo model for oxygen vacancy diffusion in praseodymium doped ceria: Applications to materials design." *Journal of Solid State Chemistry* 184:811–817. https://doi.org/10.1016/j.jssc.2011.02.004.

Dholabhai, Pratik P., Shahriar Anwar, James B. Adams, Peter A. Crozier, and Renu Sharma. 2012. "Predicting the optimal dopant concentration in gadolinium doped ceria: A kinetic lattice Monte Carlo approach." *Modelling and Simulation in Materials Science and Engineering* 20:13. https://doi.org/10.1088/0965-0393/20/1/015004.

Dickey, Elizabeth C, Xudong Fan, and Stephen J Pennycook. 2001. "Structure and Chemistry of Yttria-Stabilized Cubic-Zirconia Symmetric Tilt Grain Boundaries." *Journal of the American Ceramic Society* 84 (6): 1361–1368. https://doi.org/10.1111/j.1151-2916.2001.tb00842.x.

Diercks, David R, Jianhua Tong, Huayang Zhu, Robert Kee, George Baure, Juan C Nino, Ryan O'Hayre, and Brian P Gorman. 2016. "Three-dimensional quantification of composition and electrostatic potential at individual grain boundaries in doped ceria." *Journal of Materials Chemistry A* 4 (14): 5167–5175. https://doi.org/10.1039/C5TA10064J.

Dimroth, Frank, and Sarah Kurtz. 2007. "High-Efficiency Multijunction Solar Cells." *MRS Bulletin* 32 (3): 230–235. https://doi.org/10.1557/mrs2007.27.

Dion, Max, Henrik Rydberg, Elsebeth Schröder, David C Langreth, and Bengt I Lundqvist. 2004. "Van der Waals Density Functional for General Geometries." *Physical Review Letters* 92 (24): 246401. https://doi.org/10.1103/PhysRevLett.92.246401.

Dittmer, Anneke, Robert Izsak, Frank Neese, and Dimitrios Maganas. 2019. "Accurate Band Gap Predictions of Semiconductors in the Framework of the Similarity Transformed Equation of Motion Coupled Cluster Theory." *Inorganic Chemistry* 58 (14): 9303–9315. https://doi.org/10.1021/acs.inorgchem.9b00994.

Dobson, John F., Jun Wang, and Tim Gould. 2002. "Correlation energies of inhomogeneous many-electron systems." *Physical Review B* 66 (081108). https://doi.org/10.1103/PhysRevB.66.081108.

Dong, Renhao, Tao Zhang, and Xinliang Feng. 2018. "Interface-Assisted Synthesis of 2D Materials: Trend and Challenges." *Chem Rev* 118 (13): 6189–6235. https://doi.org/10.1021/acs.chemrev.8b00056.

Dudarev, S. L., G. A. Botton, S. Y. Savrasov, C. J. Humphreys, and A. P. Sutton. 1998. "Electron-energy-loss spectra and the structural stability of nickel oxide: An LSDA+U study." *Physical Review B* 57 (3): 1505–1509. https://doi.org/10.1103/PhysRevB.57.1505.

Eyring, L. 1979. *Handbook on the Physics and Chemistry of Rare Earths.* Edited by K. A. Gschneider and L. Eyring. 3:1–664. North-Holland, Amsterdam.

Feng, B., H. Hojo, T. Mizoguchi, H. Ohta, S. D. Findlay, Y. Sato, N. Shibata, T. Yamamoto, and Y. Ikuhara. 2012. "Atomic structure of a $\sigma 3$ [110]/(111) grain boundary in CeO_2." *Applied Physics Letters* 100:3–6. https://doi.org/10.1063/1.3682310.

Feng, B., N. R. Lugg, A. Kumamoto, Y. Ikuhara, and N. Shibata. 2017. "Direct Observation of Oxygen Vacancy Distribution across Yttria-Stabilized Zirconia Grain Boundaries." *ACS Nano* 11 (11): 11376–11382.

Feng, Bin, Issei Sugiyama, Hajime Hojo, Hiromichi Ohta, Naoya Shibata, and Yuichi Ikuhara. 2016. "Atomic structures and oxygen dynamics of CeO_2 grain boundaries." *Scientific Reports* 6:20288. https://doi.org/10.1038/srep20288.

Feng, Ya-Xin, Jia-Xiang Shang, Zeng-Hui Liu, and Guang-Hong Lu. 2015. "The energy and structure of (110) twist grain boundary in tungsten." *Applied Surface Science* 357, Part:262–267.

Figueiredo, F. M. L., and F. M. B. Marques. 2013. "Electrolytes for solid oxide fuel cells." *Wiley Interdisciplinary Reviews: Energy and Environment* 2 (1): 52–72. https://doi.org/10.1002/wene.23.

Fischer, Julia Melisande, Michelle Hunter, Marlies Hankel, Debra J. Searles, Amanda J. Parker, and Amanda S. Barnard. 2020. "Accurate prediction of binding energies for two-dimensional catalytic materials using machine learning." *ChemCatChem* 12 (20): 5109–5120. https://doi.org/10.1002/cctc.202000536.

Franceschetti, D. R. 1981. "Local Thermodynamic Formalism for Space-Charge in Ionic-Crystals." *Solid State Ionics* 2 (1): 39–42.

Frechero, M. A., M. Rocci, G. Sánchez-Santolino, Amit Kumar, J. Salafranca, Rainer Schmidt, M. R. Díaz-Guillén, et al. 2015. "Paving the way to nanoionics: Atomic origin of barriers for ionic transport through interfaces." *Scientific Reports* 5:1–9.

Fu, Szu Pei P., Rolf Ryham, Andreas Klöckner, Matt Wala, Shidong Jiang, and Yuan Nan Young. 2020. "Simulation of multiscale hydrophobic lipid dynamics via efficient integral equation methods." *Multiscale Modeling and Simulation* 18 (1): 79–103. https://doi.org/10.1137/18M1219503.

Fuchs, Martin, and Xavier Gonze. 2002. "Accurate density functionals: Approaches using the adiabatic-connection fluctuation-dissipation theorem." *Physical Review B* 65 (23): 235109. https://doi.org/10.1103/PhysRevB.65.235109.

Gao, Chao, Daxue Du, Dong Ding, Feiyang Qiao, and Wenzhong Shen. 2022. "A Review on Monolithic Perovskite/c-Si Tandem Solar Cell: Progress, Challenges, and Opportunities." *Journal of Materials Chemistry A,* https://doi.org/10.1039/D2TA01470J.

Gaye, Amie. 2008. "Access to energy and human development." *Human Development Report,* http://hdr.undp.org/sites/default/files/gaye%5C%5Famie.pdf.

Geim, A. K., and I. V. Grigorieva. 2013. "Van der Waals heterostructures." *Nature* 499 (July): 419–425. https://doi.org/10.1038/nature12385.

Genreith-Schriever, Annalena R., Jana P. Parras, Henrik J. Heelweg, and Roger A. De Souza. 2020. "The Intrinsic Structural Resistance of a Grain Boundary to Transverse Ionic Conduction." *ChemElectroChem* 7 (23): 4718–4723. https://doi.org/10.1002/celc.202000773.

Gerward, L., and J. Staun Olsen. 1993. "Powder diffraction analysis of cerium dioxide at high pressure." *Powder Diffraction* 8 (2): 127–129. https://doi.org/10.1017/S0885715600017966.

Geurts, Pierre, Damien Ernst, and Louis Wehenkel. 2006. "Extremely randomized trees." *Mach Learn* 63 (1): 3–42. https://doi.org/10.1007/s10994-006-6226-1.

Gjerding, Morten, Thorbjørn Skovhus, Asbjørn Rasmussen, Fabian Bertoldo, Ask Hjorth Larsen, Jens Jørgen Mortensen, and Kristian Sommer Thygesen. 2021. "Atomic Simulation Recipes: A Python framework and library for automated workflows." *Computational Materials Science* 199:110731. https://doi.org/10.1016/j.commatsci.2021.110731.

Goble, Carole, Sarah Cohen-Boulakia, Stian Soiland-Reyes, Daniel Garijo, Yolanda Gil, Michael R. Crusoe, Kristian Peters, and Daniel Schober. 2020. "FAIR Computational Workflows." *Data Intelligence* 2, nos. 1-2 (January): 108–121. https://doi.org/10.1162/dint_a_00033.

Gordon, Meghan, and Maya Weber. 2022. *Nickel market volatility hits China's battery industry.* Edited by Pritish Raj. https://www.spglobal.com/commodity-insights/en/market-insights/latest-news/oil/100621-global-energy-demand-to-grow-47-by-2050-with-oil-still-top-source-us-eia. [Online; accessed 26. Mar. 2022], March.

Gregori, Giuliano, Rotraut Merkle, and Joachim Maier. 2017. "Ion conduction and redistribution at grain boundaries in oxide systems." *Progress in Materials Science* 89:252–305. https://doi.org/10.1016/j.pmatsci.2017.04.009.

Grimme, Stefan, and Frank Neese. 2007. "Double-hybrid density functional theory for excited electronic states of molecules." *The Journal of chemical physics* 127 (154116). https://doi.org/10.1063/1.2772854.

Guo, Cun Xin, Jian Xin Wang, Chang Rong He, and Wei Guo Wang. 2013. "Effect of alumina on the properties of ceria and scandia co-doped zirconia for electrolyte-supported SOFC." *Ceramics International* 39, no. 8 (December): 9575–9582. https://doi.org/10.1016/j.ceramint.2013.05.076.

Guo, X., and J. Maier. 2001. "Grain boundary blocking effect in zirconia: A Schottky barrier analysis." *Journal of the Electrochemical Society* 148 (3): E121–E126.

Guo, X., W. Sigle, and J. Maier. 2003. "Blocking grain boundaries in yttria-doped and undoped ceria ceramics of high purity." *Journal of the American Ceramic Society* 86 (1): 77–87.

Gurarslan, Alper, Yifei Yu, Liqin Su, Yiling Yu, Francisco Suarez, Shanshan Yao, Yong Zhu, Mehmet Ozturk, Yong Zhang, and Linyou Cao. 2014. "Surface-energy-assisted perfect transfer of centimeter-scale monolayer and few-layer MoS_2 films onto arbitrary substrates." *ACS Nano* 8 (11): 11522–11528. https://doi.org/10.1021/nn5057673.

Gusakova, Julia, Xingli Wang, Li Lynn Shiau, Anna Krivosheeva, Victor Shaposhnikov, Victor Borisenko, Vasilii Gusakov, and Beng Kang Tay. 2017. "Electronic Properties of Bulk and Monolayer TMDs: Theoretical Study Within DFT Framework (GVJ-2eMethod)." *Physica Status Solidi A,* no. 1700218, https://doi.org/10.1002/pssa.201700218.

H., Bruus, and Flensberg K. 2002. *Introduction to Many-Body Quantum Theory in Condensed Matter Physics.* 152–183. Oxford University Press.

Haastrup, Sten, Mikkel Strange, Mohnish Pandey, Thorsten Deilmann, Per S. Schmidt, Nicki F. Hinsche, Morten N. Gjerding, et al. 2018. "The Computational 2D Materials Database: High-Throughput Modeling and Discovery of Atomically Thin Crystals." *2D Materials* 5 (4). https://doi.org/10.1088/2053-1583/aacfc1.

Hagedorn, K. E. 1973. "High-angle grain boundaries (progress in materials science, Vol. 16). H. Gleiter und B. Chalmers (1972). 274 Seit., zahlreiche Abbild. und

Tab. — Pergamon Press, Oxford, New York, Toronto, Sydney, Braunschweig, 1972.," 24:95–96. 1. https://doi.org/10.1002/maco.19730240132.

Haile, S. M., D. L. West, and J. Campbell. 1998. "The role of microstructure and processing on the proton conducting properties of gadolinium-doped barium cerate." *Journal of Materials Research* 13 (6): 1576–1595.

Haynes, W. M., and David R. Lide. 2016. "CRC Handbook of Chemistry and Physics." In *Ionic Radii in Crystals,* 97th ed. [Online; accessed 30. Mar. 2022]. Andover, England, UK: Taylor & Francis. https://doi.org/10.1201/9781315380476.

He, Lianhua, Fang Liu, Geoffroy Hautier, Micael J. T. Oliveira, Miguel A. L. Marques, Fernando D. Vila, J. J. Rehr, G.-M. Rignanese, and Aihui Zhou. 2014. "Accuracy of generalized gradient approximation functionals for density-functional perturbation theory calculations." *Physical Review B* 89 (6): 064305. https://doi.org/10.1103/PhysRevB.89.064305.

Hellman, H. 1937. *An Introduction to Quantum Chemistry.* Deuticke, Leipzig.

Henkelman, Graeme, Andri Arnaldsson, and Hannes Jønsson. 2006. "A fast and robust algorithm for Bader decomposition of charge density." *Computational Materials Science* 36. https://doi.org/10.1016/j.commatsci.2005.04.010.

Heyd, J., J.E. Peralta, G.E. Scuseria, and R.L. Martin. 2005. "Energy band gaps and lattice parameters evaluated with the Heyd-Scuseria-Ernzerhof screened hybrid functional." *Journal of Chemical Physics* 123.

Heyd, Jochen, Gustavo E. Scuseria, and Matthias Ernzerhof. 2003. "Hybrid functionals based on a screened Coulomb potential." *The Journal of Chemical Physics* 118 (18): 8207–8215. https://doi.org/10.1063/1.1564060.

———. 2006. "Erratum: "Hybrid functionals based on a screened Coulomb potential" [J. Chem. Phys. 118, 8207 (2003)]." *The Journal of Chemical Physics* 124 (21): 219906. https://doi.org/10.1063/1.2204597.

Hohenberg, P., and W. Kohn. 1964. "Inhomogeneous electron gas." *Physical review* 136 (3B): B864–B871.

Hojo, Hajime, Teruyasu Mizoguchi, Hiromichi Ohta, Scott D. Findlay, Naoya Shibata, Takahisa Yamamoto, and Yuichi Ikuhara. 2010. "Atomic structure of a CeO_2 grain boundary: The role of oxygen vacancies." *Nano Letters* 10:4668–4672. https://doi.org/10.1021/nl1029336.

Homer, Eric R., Srikanth Patala, and Jonathan L. Priedeman. 2015. "Grain Boundary Plane Orientation Fundamental Zones and Structure-Property Relationships." *Scientific reports* 5, no. 15476 (October): 1–13. https://doi.org/10.1038/srep15476.

Hou, Yusheng, Feng Xue, Liang Qiu, Zhe Wang, and Ruqian Wu. 2022. "Multifunctional two-dimensional van der Waals Janus magnet Cr-based dichalcogenide halides." *npj Computational Materials* 8 (120): 1–6. https://doi.org/10.1038/s41524-022-00802-x.

Hwang, Jin Ha, and Thomas O. Mason. 1998. "Defect chemistry and transport properties of nanocrystalline cerium oxide." *Zeitschrift für Physikalische Chemie* 207 (1-2): 21–38. https://doi.org/10.1524/zpch.1998.207.part_1_2.021.

Idrees, M., H. U. Din, Shafiq Ur Rehman, M. Shafiq, Yasir Saeed, H. D. Bui, Chuong V. Nguyen, and Bin Amin. 2020. "Electronic properties and enhanced photocatalytic performance of van der Waals heterostructures of ZnO and Janus transition metal dichalcogenides." *Physical Chemistry Chemical Physics* 22 (18): 10351–10359. https://doi.org/10.1039/d0cp01264e.

Idrees, M., M. Fawad, M. Bilal, Y. Saeed, C. Nguyen, and Bin Amin. 2020. "Van der Waals heterostructures of SiC and Janus MSSe (M = Mo, W) monolayers: A first principles study." *RSC Advances* 10 (43): 25801–25807. https://doi.org/10.1039/d0ra04433d.

Electricity Market Report–July. 2021. https://www.iea.org/reports/electricity-market-report-july-2021. [Online; accessed 26. Mar. 2022]. Paris, July.

Jain, Anubhav, Shyue Ping Ong, Wei Chen, Bharat Medasani, Xiaohui Qu, Michael Kocher, Miriam Brafman, et al. 2015. "FireWorks: a dynamic workflow system designed for high-throughput applications." *Concurrency and Computation: Practice and Experience* 27 (17): 5037–5059. https://doi.org/10.1002/cpe.3505.

Jain, Anubhav, Shyue Ping Ong, Geoffroy Hautier, Wei Chen, William Davidson Richards, Stephen Dacek, Shreyas Cholia, et al. 2013. "The Materials Project: A materials genome approach to accelerating materials innovation." *APL Materials* 1 (1): 011002. https://doi.org/10.1063/1.4812323.

Jena, Debdeep, and Aniruddha Konar. 2007. "Enhancement of carrier mobility in semiconductor nanostructures by dielectric engineering." *Physical Review Letters* 98 (13). https://doi.org/10.1103/PhysRevLett.98.136805.

Jiang, Yong, James B. Adams, Mark van Schilfgaarde, Renu Sharma, and Peter A. Crozier. 2005. "Theoretical study of environmental dependence of oxygen vacancy formation in CeO_2." *Applied Physics Letters* 87:141917.

Jie, JianShu, MouYi Weng, ShunNing Li, Dong Chen, ShuCheng Li, WeiJi Xiao, JiaXin Zheng, Feng Pan, and LinWang Wang. 2019. "A new MaterialGo database and its comparison with other high-throughput electronic structure databases for their predicted energy band gaps." *Science China Technological Sciences* 62 (8): 1423–1430. https://doi.org/10.1007/s11431-019-9514-5.

Kang, Jun, Jingbo Li, Shu-Shen Li, Jian-Bai Xia, and Lin-Wang Wang. 2013. "Electronic Structural Moire Pattern Effects on MoS_2/$MoSe_2$ 2D Heterostructure." *Nano Letters* 12:5485–5490. https://doi.org/10.1021/nl4030648.

Kim, Kyoo, Sooran Kim, J. S. Kim, Heejung Kim, J. H. Park, and B. I. Min. 2018. "Importance of the van Hove singularity in superconducting $PdTe_2$." *Physical Review B* 97 (16): 165102. https://doi.org/10.1103/PhysRevB.97.165102.

Kim, S. 2016. "Isn't the space-charge potential in ceria-based solid electrolytes largely overestimated?" *Physical Chemistry Chemical Physics* 18 (29).

Kim, S., S. K. Kim, S. Khodorov, J. Maier, and I. Lubomirsky. 2016. "On determining the height of the potential barrier at grain boundaries in ion-conducting oxides." *Physical Chemistry Chemical Physics* 18 (4): 3023–3031.

Kirklin, Scott, James E. Saal, Bryce Meredig, Alex Thompson, Jeff W. Doak, Muratahan Aykol, Stephan Rühl, and Chris Wolverton. 2015. "The Open Quantum Materials Database (OQMD): assessing the accuracy of DFT formation energies." *npj Computational Materials* 1 (1): 15010. https://doi.org/10.1038/npjcompumats.2015.10.

Klimeš, Jiři, David R Bowler, and Angelos Michaelides. 2009. "Chemical accuracy for the van der Waals density functional." *Journal of Physics: Condensed Matter* 22, no. 2 (December): 022201. https://doi.org/10.1088/0953-8984/22/2/022201.

Koettgen, Julius, Steffen Grieshammer, Philipp Hein, Benjamin O. H. Grope, Masanobu Nakayama, and Manfred Martin. 2018. "Understanding the ionic conductivity maximum in doped ceria: trapping and blocking." *Physical Chemistry Chemical Physics* 20 (21): 14291–14321. https://doi.org/10.1039/C7CP08535D.

Koettgen, Julius, and Manfred Martin. 2019. "The effect of jump attempt frequency on the ionic conductivity of doped ceria." *Journal of Physical Chemistry C* 123:19437–19446.

Koettgen, Julius, and Manfred Martin. 2020. "Infinite dilution in doped ceria and high activation energies." *Solid State Communications,* 113939.

Kohn, W., and L. J. Sham. 1965. "Self-consistent equations including exchange and correlation effects." *Physical Review* 140 (4A): A1133–A1138.

Kong, Desheng, Haotian Wang, Judy J. Cha, Mauro Pasta, Kristie J. Koski, Jie Yao, and Yi Cui. 2013. "Synthesis of MoS_2 and $MoSe_2$ films with vertically aligned layers." *Nano Letters* 13 (3): 1341–1347. https://doi.org/10.1021/nl400258t.

Kresse, G., and J. Furthmüller. 1996a. "Efficiency of ab-initio total energy calculations for metals and semiconductors using a plane-wave set." *Computational Materials Science* 6:15. https://doi.org/10.1016/0927-0256(96)00008-0.

———. 1996b. "Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set." *Physical Review B* 54:11169–11186. https://doi.org/10.1103/PhysRevB.54.11169.

Kresse, G., and J. Hafner. 1993. "Ab initio molecular dynamics for liquid metals." *Physical Review B* 47:558. https://doi.org/10.1016/0022-3093(95)00355-X.

———. 1994. "Ab initio molecular-dynamics simulation of the liquid-metal-amorphous-semiconductor transition in germanium." *Physical Review B* 49 (20): 14251. https://doi.org/10.1103/physrevb.49.14251.

Kresse, G., and D. Joubert. 1999. "From ultrasoft pseudopotentials to the projector augmented-wave method." *Physical Review B* 59:1758.

Kresse, G., M. Marsman, and J Furthmuller. 2018. *VASP Manual.* Vienna: Universitat Wien.

Krukau, A. V., O. A. Vydrov, A. F. Izmaylov, and G. E. Scuseria. 2006. "Influence of the exchange screening parameter on the performance of screened hybrid functionals." *The Journal of chemical physics* 125 (224106). https://doi.org/10.1063/1.2404663.

Kwak, No Woo, and Woo Chul Jung. 2016. "Analysis of the grain boundary conductivity of singly and doubly doped CeO_2 thin films at elevated temperature." *Acta Materialia* 108:271–278. https://doi.org/10.1016/j.actamat.2016.02.032.

Lee, H. B., F. B. Prinz, and W. Cai. 2013. "Atomistic simulations of grain boundary segregation in nanocrystalline yttria-stabilized zirconia and gadolinia-doped ceria solid oxide electrolytes." *Acta Materialia* 61 (10): 3872–3887.

Lee, Yi-Hsien, Lili Yu, Han Wang, Wenjing Fang, Xi Ling, Yumeng Shi, Cheng-Te Lin, et al. 2013. "Synthesis and Transfer of Single-Layer Transition Metal Disulfides on Diverse Surfaces." *Nano Letters* 13 (4): 1852–1857. https://doi.org/10.1021/nl400687n.

Lee, J. S., and D. Y. Kim. 2001. "Space-charge concepts on grain boundary impedance of a high-purity yttria-stabilized tetragonal zirconia polycrystal." *Journal of Materials Research* 16 (9): 2739–2751.

Lee, Kang Taek, and Eric D. Wachsman. 2014. "Role of nanostructures on SOFC performance at reduced temperatures." *MRS Bulletin* 39:783–791. https://doi.org/10.1557/mrs.2014.193.

Lee, W., H. J. Jung, M. H. Lee, Y. B. Kim, J. S. Park, R. Sinclair, and F. B. Prinz. 2012. "Oxygen Surface Exchange at Grain Boundaries of Oxide Ion Conductors." *Advanced Functional Materials* 22 (5).

Lee, Yueh-Lin, Yuhua Duan, Dane Morgan, Dan C. Sorescu, Harry Abernathy, and Gregory Hackett. 2017. "Density-Functional-Theory Modeling of Cation Diffusion in Bulk $La_{1-x}Sr_xMnO_{3\pm\delta}$ ($x = 0.0 - -0.25$) for Solid-Oxide Fuel-Cell Cathodes." *Physical Review Applied* 8 (4): 044001. https://doi.org/10.1103/PhysRevApplied.8.044001.

Lejček, Pavel. 2010. "Grain Boundaries: Description, Structure and Thermodynamics." Chap. 2 in *Grain Boundary Segregation In Metals,* 5–24. Berlin, Germany: Springer. https://doi.org/10.1007/978-3-642-12505-8.

Lesar, Richard. 2013. *Introduction to Computational Materials Science: Fundamentals to Applications.* 414. Cambridge University Press.

Li, Y., C. Gao, R. Long, and Y. Xiong. 2019. "Photocatalyst design based on two-dimensional materials." *Materials Today Chemistry* 11:197–216. https://doi.org/10.1016/j.mtchem.2018.11.002.

Lia, F., T. Ohkubo, Y.M. Chen, M. Kodzuka, F. Ye, D.R. Ou, T. Mori, and K. Hono. 2010. "Laser-Assisted Three-Dimensional Atom Probe Analysis of Dopant Distribution in Gd-Doped CeO_2." *Scripta Materialia* 63 (3): 332–335. https://doi.org/10.1016/j.scriptamat.2010.04.029.

Lin, Y., S. Fang, D. Su, K. S. Brinkman, and F. Chen. 2015. "Enhancing grain boundary ionic conductivity in mixed ionic-electronic conductors." *Nature Communications* 6.

Liu, Li, Tie Chen, and Yanhong Yin. 2016. "Energy Consumption and Quality of Life: Energy Efficiency Index." *Energy Procedia* 88:224–229. https://doi.org/10.1016/j.egypro.2016.06.152.

Liu, Yue, Tianlu Zhao, Wangwei Ju, and Siqi Shi. 2017. "Materials discovery and design using machine learning." *Journal of Materiomics* 3 (3): 159–177. https://doi.org/10.1016/j.jmat.2017.08.002.

Lu, Ang Yu, Hanyu Zhu, Jun Xiao, Chih Piao Chuu, Yimo Han, Ming Hui Chiu, Chia Chin Cheng, et al. 2017. "Janus monolayers of transition metal dichalcogenides." *Nature Nanotechnology* 12 (8): 744–749. https://doi.org/10.1038/nnano.2017.100.

Luo, Bin, Gang Liu, and Lianzhou Wang. 2016. "Recent advances in 2D materials for photocatalysis." *Nanoscale* 8:6904–6920. https://doi.org/10.1039/c6nr00546b.

Ma, Xiangchao, Yi Tian, Pei Zhao, Xin Wu, Tao Jing, and Jianqi Zhang. 2022. "Janus MoCrSSe monolayer: A strong two dimensional polar antiferromagnet." *Applied Surface Science* 581 (April). https://doi.org/10.1016/j.apsusc.2022.152420.

Maghirang, Aniceto B., Zhi Quan Huang, Rovi Angelo B. Villaos, Chia Hsiu Hsu, Liang Ying Feng, Emmanuel Florido, Hsin Lin, Arun Bansil, and Feng Chuan Chuang. 2019. "Predicting two-dimensional topological phases in Janus materials by substitutional doping in transition metal dichalcogenide monolayers." *npj 2D Materials and Applications* 3, no. 35 (1). https://doi.org/10.1038/s41699-019-0118-2.

Maier, J. 1995. "Ionic-conduction in Space Charge Regions." *Progress in Solid State Chemistry* 23 (3): 171–263.

Mak, Kin Fai, Changgu Lee, James Hone, Jie Shan, and Tony F Heinz. 2010. "Atomically thin MoS_2: A new direct-gap semiconductor." *Physical Review Letters* 105 (13): 136805. https://doi.org/10.1103/PhysRevLett.105.136805.

Mathew, Kiran, Joseph H. Montoya, Alireza Faghaninia, Shyam Dwarakanath, Muratahan Aykol, Hanmei Tang, Iek-heng Chu, et al. 2017. "Atomate: A high–level interface to generate, execute, and analyze computational materials science workflows." *Computational Material Science* 139:140–152. https://doi.org/10.1016/j.commatsci.2017.07.030.

Mathew, Kiran, Arunima K Singh, Joshua J Gabriel, Kamal Choudhary, Susan B Sinnott, Albert V Davydov, Francesca Tavazza, and Richard G Hennig. 2016. "MPInterfaces: A Materials Project based Python tool for high-throughput com-

putational screening of interfacial systems." *Computational Materials Science* 122:183–190. https://doi.org/10.1016/j.commatsci.2016.05.020.

Mazur, Allan. 2011. "Does increasing energy or electricity consumption improve quality of life in industrial nations?" *Energy Policy* 39 (5): 2568–2572. https://doi.org/10.1016/j.enpol.2011.02.024.

Mebane, D. S., and R. A. De Souza. 2015. "A generalised space-charge theory for extended defects in oxygen-ion conducting electrolytes: from dilute to concentrated solid solutions." *Energy & Environmental Science* 8 (10): 2935–2940.

Minervini, Licia, Matthew O. Zacate, and Robin W. Grimes. 1999. "Defect cluster formation in M_2O_3-doped CeO_2." *Solid State Ionics* 116:339–349.

Moradpoor, I., and M. Ebrahimi. 2016. "Combined solid oxide fuel cell, micro-gas turbine and orgainic Rankine cycle for power generation (SOFC-MGT-ORC)." *Energy Conversion and Management* 116:120–133. https://doi.org/10.1016/j.enconman.2016.02.080.

Mounet, N., M. Gibertini, P. Schwaller, D. Campi, A. Merkys, A. Marrazzo, T. Sohier, et al. 2018. "Two-dimensional materials from high-throughput computational exfoliation of experimentally known compounds." *Nature Nanotechnology* 13 (3): 246. https://doi.org/10.1038/s41565-017-0035-5.

Muench, W. von, C. Gessert, and M. E. Koeniger. 1976. "Photodiodes and junction field-effect transistors with high UV sensitivity." *IEEE Trans Electron Devices* 23, no. 11 (November): 1203–1207. https://doi.org/10.1109/T-ED.1976.18579.

Nafsin, N., T. J. A. Aguiar, A. M. Aoki, K. van Benthem Thron, and R. H. R. Castro. 2017. "Thermodynamics versus kinetics of grain growth control in nanocrystalline zirconia." *Acta Materialia* (136): 224–234.

Nakajima, A., A. Yoshihara, and M. Ishigame. 1994. "Defect-induced Raman spectra in doped CeO_2." *Physical Review B* 50 (18): 13297–13307. https://doi.org/10.1103/PhysRevB.50.13297.

Nolan, Michael, and Graeme W. Watson. 2005. "The electronic structure of alkali doped alkaline earth metal oxides: Li doping of MgO studied with DFT-GGA and GGA+U." *Surface Science* 586 (1-3): 25–37. https://doi.org/10.1016/j.susc.2005.04.046.

Novoselov, K. S. 2011. "Nobel Lecture: Graphene: Materials in the Flatland." *Review of Modern Physics* 83 (3): 837–849. https://doi.org/10.1103/RevModPhys.83.837.

Novoselov, K. S., V. I. Falko, L. Colombo, P. R. Gellert, M. G. Schwab, and K. Kim. 2012. "A roadmap for graphene." *Nature Review* 490:192–200. https://doi.org/10.1038/nature11458.

Novoselov, K. S., A. K. Geim, S. V. Morozov, D. Jiang, Y. Zhang, S. V. Dubonos, I. V. Grigorieva, and A. A. Firsov. 2004. "Electric Field Effect in Atomically Thin Carbon Films." *Science* 306 (5696): 666–669. https://doi.org/10.1126/science.1102896.

Novoselov, K. S., A. Mishchenko, A. Carvalho, and A. H. Castro Neto. 2016. "2D materials and van der Waals heterostructures." *Science* 353. https://doi.org/10.1126/science.aac9439.

Oil Spills. 2017. [Online; accessed 26. Mar. 2022], January. https://medlineplus.gov/oilspills.html.

Oil spills. 2020. [Online; accessed 26. Mar. 2022], September. https://www.noaa.gov/education/resource-collections/ocean-coasts/oil-spills.

Ong, Shyue Ping, William Davidson Richards, Anubhav Jain, Geoffroy Hautier, Michael Kocher, Shreyas Cholia, Dan Gunter, Vincent L. Chevrier, Kristin A. Persson, and Gerbrand Ceder. 2013. "Python Materials Genomics (pymatgen): A robust, open-source python library for materials analysis." *Computational Materials Science* 68:314–319. https://doi.org/10.1016/j.commatsci.2012.10.028.

Orlovskaya, Nina, and Nigel Browning. 2004. "Mixed Ionic Electronic Conducting Perovskites for Advanced Energy Systems." In *Vacancy Segregation at Grain Boundaries in Ceramic Oxides,* vol. 173. The Netherlands, Kluwer: Springer Science & Business Media.

Ouchi, Hideki, Takashi Mukai, Takafumi Kamei, and Masahiro Okamura. 1979. "Silicon p-n junction photodiodes sensitive to ultraviolet radiation." *IEEE Transactions on Electron Devices* 26:1965–1969. https://doi.org/10.1109/T-ED.1979.19803.

Global primary energy consumption by source. 2022. https://ourworldindata.org/grapher/global-energy-consumption-source?country=~OWID%5FWRL. [Online; accessed 5. Jul. 2022].

Pablo, Juan J. de, Nicholas E. Jackson, Michael A. Webb, Long-Qing Chen, Joel E. Moore, Dane Morgan, Ryan Jacobs, et al. 2019. "New frontiers for the materials genome initiative." *npj Comput Mater* 5 (41): 1–23. https://doi.org/10.1038/s41524-019-0173-4.

Paier, J., C. Penschke, and J. Sauer. 2013. "Oxygen Defects and Surface Chemistry of Ceria: Quantum Chemical Studies Compared to Experiment." *Chemical Reviews* 113 (6): 3949–3985.

Paul, Joshua T., Arunima K. Singh, Z. Dong, H. Zhuang, B. C. Revard, B. Rijal, M. Ashton, et al. 2017. "Computational methods for 2D materials: discovery, property characterization, and application design." *Journal of Physics: Condensed Matter* 29 (47): 473001. https://doi.org/10.1088/1361-648X/aa9305.

Pedregosa, Fabian, Gaël Varoquaux, Alexandre Gramfort, Bertrand Thirion, Olivier Grisel, Peter Prettenhofer, Ron Weiss, et al. 2011. "Scikit-learn: Machine Learning in Python." *Journal of Machine Learning Research* 12:2825–2830. http://scikit-learn.sourceforge.net..

Perdew, J. P. 1985. *Physical review Lett.* 55:1665.

Perdew, J. P., K. Burke, and M. Ernzerhof. 1996. "Generalized Gradient Approximation Made Simple." *Physical Review Letters* 77:3865.

———. 1997. "Erratum Generalized Gradient Approximation Made Simple." *Physical review Lett.* 78:1396.

Perdew, J.P. 1985. "Density functional theory and the band gap problem." *Journal of Quantum Chemistry* 28:497–523. https://doi.org/10.1002/qua.560280846.

Perdew, John P., Adrienn Ruzsinszky, Jianmin Tao, Viktor N. Staroverov, Gustavo E. Scuseria, and Gábor I. Csonka. 2005. "Prescription for the design and selection of density functional approximations: More constraint satisfaction with fewer fits." *The Journal of chemical physics* 123. https://doi.org/10.1063/1.1904565.

Pleshakov, Edward. 2008. *File:Crystallite.jpg - Wikimedia Commons.* https://commons.wikimedia.org/wiki/File:Crystallite.jpg. [Online; accessed 6. Jul. 2022].

Primary energy consumption by world region. 2022. https://ourworldindata.org/grapher/primary-energy-consumption-by-region. [Online; accessed 26. Mar. 2022], March.

Qin, Ying, Mohammed Sayyad, Alejandro R.P. Montblanch, Matthew S.G. Feuer, Dibyendu Dey, Mark Blei, Renee Sailus, et al. 2022. "Reaching the Excitonic Limit in 2D Janus Monolayers by In Situ Deterministic Growth." *Advanced Materials* 34 (6). https://doi.org/10.1002/adma.202106222.

Randle, V. 1993. *The Measurement of Grain Boundary Geometry.* 1–169. Institute of Physics Publishing.

Revard, Benjamin C., William W. Tipton, Anna Yesypenko, and Richard G. Hennig. 2016. "Grand-canonical evolutionary algorithm for the prediction of two-dimensional materials." *Physical Review B* 93 (5): 054117. https://doi.org/10.1103/PhysRevB.93.054117.

Rhodes, Daniel, Sang Hoon Chae, Rebeca Ribeiro-Palau, and James Hone. 2019. "Disorder in van der Waals heterostructures of 2D materials." *Nature materials* 18:541–549. https://doi.org/10.1038/s41563-019-0366-8.

Rhone, Trevor David, Wei Chen, Shaan Desai, Steven B. Torrisi, Daniel T. Larson, Amir Yacoby, and Efthimios Kaxiras. 2020. "Data-driven studies of magnetic two-dimensional materials." *Scientific reports* 10, no. 15795 (September): 1–11. https://doi.org/10.1038/s41598-020-72811-z.

Riis-Jensen, Anders C., Thorsten Deilmann, Thomas Olsen, and Kristian S. Thygesen. 2019. "Classifying the Electronic and Optical Properties of Janus Monolayers." *ACS Nano* 13 (11): 13354–13364. https://doi.org/10.1021/acsnano.9b06698.

Román-Pérez, G., and J. M. Soler. 2009. *Physical review Lett.* 103.

Roser, Max. 2013. "Economic Growth." Https://ourworldindata.org/economic-growth, *Our World in Data.*

Roser, Max, and Esteban Ortiz-Ospina. 2013. "Global Extreme Poverty." Https://ourworldindata.org/extreme-poverty, *Our World in Data.*

Roser, Max, Esteban Ortiz-Ospina, and Hannah Ritchie. 2013. "Life Expectancy." Https://ourworldindata.org/life-expectancy, *Our World in Data.*

Rupp, Jennifer L. M., Anna Infortuna, and Ludwig J. Gauckler. 2006. "Microstrain and self-limited grain growth in nanocrystalline ceria ceramics." *Acta Materialia* 54, no. 7 (April): 1721–1730. https://doi.org/10.1016/j.actamat.2005.11.032.

Rydberg, H., M. Dion, N. Jacobson, E. Schröder, P. Hyldgaard, S. I. Simak, D. C. Langreth, and B. I. Lundqvist. 2003. "Van der Waals Density Functional for Layered Structures." *Physical Review Letters* 91 (126402). https://doi.org/10.1103/PhysRevLett.91.126402.

Saito, Yu, Tsutomu Nojima, and Yoshihiro Iwasa. 2016. "Highly crystalline 2D superconductors." *Nature Reviews Materials* 2 (16094): 1–18. https://doi.org/10.1038/natrevmats.2016.94.

Sato, Kiminori. 2015. "Grain-boundary structures associated with ionic transport in Gd-doped ceria nanostructured electrolyte." *Journal of Physical Chemistry C* 119 (10): 5734–5738. https://doi.org/10.1021/acs.jpcc.5b00155.

Sayle, Thi X.T., Stephen C. Parker, and Dean C. Sayle. 2006. "Ionic conductivity in nano-scale CeO_2/YSZ heterolayers." *Journal of Materials Chemistry* 16 (11): 1067–1081. https://doi.org/10.1039/b511547g.

Schmitt, Rafael, Andreas Nenning, Olga Kraynis, Roman Korobko, Anatoly I. Frenkel, Igor Lubomirsky, Sossina M. Haile, and Jennifer L.M. Rupp. 2020. "A review of defect structure and chemistry in ceria and its solid solutions." *Chemical Society Reviews* 49 (2): 554–592. https://doi.org/10.1039/c9cs00588a.

Schweiger, S., R. Pfenninger, W. J. Bowman, U. Aschauer, and J. L. M. Rupp. 2017. "Designing Strained Interface Heterostructures for Memristive Devices." *Advanced Materials* 29 (15): 11.

Schweinfest, R., Anthony Paxton, and M. W. Finnis. 2004. "Bismuth embrittlement of copper is an atomic size effect." *Nature* 432, no. 7020 (December): 1008–1011. https://doi.org/10.1038/nature03198.

Setyawan, W., R.M. Gaume, S. Lam; R. Feigelson, and S. Curtarolo. 2011. "High-throughput combinatorial database of electronic band structures for inorganic scintillator materials." *ACS Combinatorial Science.*

Shi, Huangang, Chao Su, Ran Ran, Jiafeng Cao, and Zongping Shao. 2020. "Electrolyte materials for intermediate-temperature solid oxide fuel cells." *Progress in Natural Science: Materials International* 30 (6): 764–774. https://doi.org/10.1016/j.pnsc.2020.09.003.

Shi, Zhe, He Zhang, Karim Khan, Rui Cao, Ye Zhang, Chunyang Ma, Ayesha Khan Tareen, Yuanfei Jiang, Mingxing Jin, and Han Zhang. 2022. "Two-dimensional materials toward Terahertz optoelectronic device applications." *Journal of Photochemistry and Photobiology C: Photochemistry Reviews* 51:100473. https://doi.org/10.1016/j.jphotochemrev.2021.100473.

Shibata, Naoya, Fumiyasu Oba, Takahisa Yamamoto, and Yuichi Ikuhara. 2004. "Structure, energy and solute segregation behaviour of [110] symmetric tilt grain

boundaries in yttria-stabilized cubic zirconia." *Philosophical Magazine* 84:2381–2415.

Shibata, Naoya, Fumiyasu Oba, Takahisa Yamamoto, Yuichi Ikuhara, and Taketo Sakuma. 2002. "Atomic structure and solute segregation of a $\Sigma = 3$, [110]/{111} grain boundary in an yttria-stabilized cubic zirconia bicrystal." *Philosophical Magazine Letters* 82:393–400. https://doi.org/10.1080/09500830210137407.

Shibata, Naoya, Fumiyasu Oba, Takahisa Yamamoto, Taketo Sakuma, and Yuichi Ikuhara. 2003. "Grain-boundary faceting at a $\Sigma = 3$, [110]/{112} grain boundary in a cubic zirconia bicrystal." *Philosophical Magazine* 83:2221–2246. https://doi.org/10.1080/0141861031000108231.

Shirpour, M., B. Rahmati, W. Sigle, P. A. van Aken, R. Merkle, and J. Maier. 2012. "Dopant Segregation and Space Charge Effects in Proton-Conducting $BaZrO_3$ Perovskites." *Journal of Physical Chemistry C* 116 (3): 2453–2461.

Singh, Arunima K., and Richard G. Hennig. 2014a. "Computational prediction of two-dimensional group-IV mono-chalcogenides." *Applied Physics Letters* 105 (4): 042103. https://doi.org/10.1063/1.4891230.

———. 2014b. "Computational synthesis of single-layer GaN on refractory materials." *Applied Physics Letters* 105 (5). https://doi.org/10.1063/1.4892351.

Singh, Arunima K., Richard G. Hennig, Albert V. Davydov, and Francesca Tavazza. 2015. "Al_2O_3 as a suitable substrate and a dielectric layer for n-layer MoS_2." *Applied Physics Letters* 107 (5): 053106. https://doi.org/10.1063/1.4928179.

Singh, Arunima K., Kiran Mathew, Houlong L. Zhuang, and Richard G. Hennig. 2015. "Computational screening of 2D materials for photocatalysis." *Journal of Physical Chemistry Letters* 6 (6): 1087–1098. https://doi.org/10.1021/jz502646d.

Singh, Arunima K., Houlong L. Zhuang, and Richard G. Hennig. 2014. "Ab initio synthesis of single-layer III-V materials." *Physical Review B - Condensed Matter and Materials Physics* 89:1–10. https://doi.org/10.1103/PhysRevB.89.245431.

Singhal, S. C., and K. E. Kendall. 2003. "High Temperature Solid Oxide Fuel Cells: Fundamentals, Design and Applications." New York: Elsevier.

Siriwardane, Edirisuriya M.D., Rajendra P. Joshi, Neeraj Kumar, and Deniz Çaklr. 2020. "Revealing the Formation Energy-Exfoliation Energy-Structure Correlation of MAB Phases Using Machine Learning and DFT." *ACS Applied Materials and Interfaces* 12 (26): 29424–29431. https://doi.org/10.1021/acsami.0c03536.

Smil, Vaclav. 2017. *Energy Transitions: Global and National Perspectives.* [Online; accessed 26. Mar. 2022]. http://vaclavsmil.com/2016/12/14/energy-transitions-global-and-national-perspectives-second-expanded-and-updated-edition/.

Sõderlind, Per, G. Kotliar, K. Haule, P. M. Oppeneer, and D. Guillaumont. 2010. "Computational modeling of actinide materials and complexes." *MRS Bulletin* 35:883–888.

Sorkun, Murat Cihan, Séverin Astruc, J. M. Vianney A. Koelman, and Er Süleyman. 2020. *V2DB: Virtual 2D Materials Database.* V. V1. https://doi.org/10.7910/DVN/SNCZF4.

Standards, National Institute of, and Technology. *NIST Inorganic Crystal Structure Database.* NIST Standard Reference Database Number 3. Gaithersburg MD. https://doi.org/10.18434/M32147.

Stephens, P. J., F. J. Devlin, C. F. Chabalowski, and M. J. Frisch. 1994. "Ab Initio Calculation of Vibrational Absorption and Circular Dichroism Spectra Using Density Functional Force Fields." *The Journal of Physical Chemistry* 98 (45): 11623–11627. https://doi.org/10.1021/j100096a001.

Sun, Qian, Zhaoming Fu, and Zongxian Yang. 2018. "Effects of rare-earth doping on the ionic conduction of $CeO_2 insolidoxidefuelcells$." *Ceramics International* 44 (4): 3707–3711. https://doi.org/10.1016/j.ceramint.2017.11.149.

Sunarso, J., S. Baumann, J. M. Serra, W. A. Meulenberg, S. Liu, Y. S. Lind, and J. C. Diniz da Costa. 2008. "Mixed ionic–electronic conducting (MIEC) ceramic-based membranes for oxygen separation." *Journal of Membrane Science* 320:13–41.

Sutton, A. P., and R. W. Balluffi. 1995. "Interfaces in Crystalline Materials," 1–798. Clarendon Press.

Talaei, M. S., N. Nouri, and S. Ziaei-Rad. 2019. "An optimized approach for computing coincidence-site-lattice grain boundary energy." *Computational Condensed Matter* 19 (June): e00363. https://doi.org/10.1016/j.cocom.2019.e00363.

Tang, Lei, Junyang Tan, Huiyu Nong, Bilu Liu, and Hui-Ming Cheng. 2021. "Chemical Vapor Deposition Growth of Two-Dimensional Compound Materials: Controllability, Material Quality, and Growth Mechanism." *Accounts of Materials Research* 2 (1): 36–47. https://doi.org/10.1021/accountsmr.0c00063.

Tang, W., E. Sanville, and G. Henkelman. 2009. "A grid-based Bader analysis algorithm without lattice bias." *Journal of Physics Condensed Matter* 21 (8). https://doi.org/10.1088/0953-8984/21/8/084204.

Tao, Jianmin, John P. Perdew, Viktor N. Staroverov, and Gustavo E. Scuseria. 2003. "Climbing the Density Functional Ladder: Nonempirical Meta–Generalized Gradient Approximation Designed for Molecules and Solids." *Physical Review Letters* 91 (14): 146401. https://doi.org/10.1103/PhysRevLett.91.146401.

Tolba, Sarah A., Kareem M. Gameel, Basant A. Ali, Hossam A. Almossalami, and Nageh K. Allam. 2018. *Recent Progresses of Theory and Application: The DFT+U: Approaches, Accuracy, and Applications, Density Functional Calculations.* IntechOpen. https://doi.org/10.5772/intechopen.72020.

Torrisi, Steven, Arunima K. Singh, Joseph H. Montoya, Tathagata Biswas, and Kristin A. Persson. 2020. "Two-Dimensional Forms of Robust CO_2 Reduction Photocatalysts." *npj 2D Materials and Applications* 4 (1): 1–10. https://doi.org/10.1038/s41699-020-0154-y.

Torrisi, Steven B., Matthew R. Carbone, Brian A. Rohr, Joseph H. Montoya, Yang Ha, Junko Yano, Santosh K. Suram, and Linda Hung. 2020. "Random forest machine learning models for interpretable X-ray absorption near-edge structure spectrum-property relationships." *npj Computational Materials* 6 (109): 1–11. https://doi.org/10.1038/s41524-020-00376-6.

Trovarelli, A. 2002. *Catalysis by ceria and related materials.* London: Imperial College Press.

Tuller, H. L. 2006. "Ionic Conduction and Applications: Springer Handbook of Electronic and Photonic Materials," 213–228. New York: Springer Verlag.

Tuller, H. L., and S. R. Bishop. 2010. "Tailoring Material Properties through Defect Engineering." *Chemistry Letters* 39 (12): 1226–1231.

Tuller, Harry L., and Sean R. Bishop. 2011. "Point Defects in Oxides: Tailoring Materials Through Defect Engineering." *Annual Review of Materials Research* 41:369–398. https://doi.org/10.1146/annurev-matsci-062910-100442.

Tusche, C., H. L. Meyerheim, and J. Kirschner. 2007. "Observation of depolarized ZnO (0001) monolayers: Formation of unreconstructed planar sheets." *Physical Review Letters* 99 (2). https://doi.org/10.1103/PhysRevLett.99.026102.

Uesugi, Tokuteru, and Kenji Higashi. 2011. "First-principles calculation of grain boundary energy and grain boundary excess free volume in aluminum: Role of grain boundary elastic energy." *Journal of Materials Science* 46 (12): 4199–4205. https://doi.org/10.1007/s10853-011-5305-2.

Uhlmann, David M. 2020. *BP paid a steep price for the Gulf oil spill but for the US a decade later, it's business as usual.* [Online; accessed 26. Mar. 2022], April. https://theconversation.com/bp-paid-a-steep-price-for-the-gulf-oil-spill-but-for-the-us-a-decade-later-its-business-as-usual-136905.

Vanpoucke, Danny E. P., Patrick Bultinck, Stefaan Cottenier, Veronique Van Speybroeckb, and Isabel Van Driessche. 2014. "Aliovalent doping of CeO_2: DFT study of oxidation state and vacancy effects." *Journal of Materials Chemistry A* 2:13723. https://doi.org/10.1039/c4ta02449d.

Vivanco-Benavides, Luis Enrique, Claudia Lizbeth Martínez-González, Cecilia Mercado-Zúñiga, and Carlos Torres-Torres. 2022. "Machine learning and materials informatics approaches in the analysis of physical properties of carbon nanotubes: A review." *Computational Materials Science* 201 (January). https://doi.org/10.1016/j.commatsci.2021.110939.

Vydrov, Oleg A., and Gustavo E. Scuseria. 2004. "Effect of the Perdew–Zunger self-interaction correction on the thermochemical performance of approximate density functionals." *The Journal of Chemical Physics* 121 (17): 8187–8193. https://doi.org/10.1063/1.1794633.

Wachsman, E. D., and K. T. Lee. 2011. "Lowering the Temperature of Solid Oxide Fuel Cells." *Science* 334 (6058): 935–939.

Wachsman, E. D., C. A. Marlowe, and K. T. Lee. 2012. "Role of solid oxide fuel cells in a balanced energy strategy." *Energy & Environmental Science* 5 (2): 5498–5509.

Walsh, Lee A., and Christopher L. Hinkle. 2017. "van der Waals epitaxy: 2D materials and topological insulators." *Applied Materials Today* 9:504–515. https://doi.org/10.1016/j.apmt.2017.09.010.

Ward, Logan, Alexander Dunn, Alireza Faghaninia, Nils E.R. Zimmermann, Saurabh Bajaj, Qi Wang, Joseph Montoya, et al. 2018. "Matminer: An open source toolkit for materials data mining." *Computational Materials Science* 152:60–69. https://doi.org/10.1016/j.commatsci.2018.05.018.

Ward, Logan, Ruoqian Liu, Amar Krishna, Vinay I. Hegde, Ankit Agrawal, Alok Choudhary, and Chris Wolverton. 2017. "Including crystal structure attributes in

machine learning models of formation energies via Voronoi tessellations." *Physical Review B* 96 (2): 024104. https://doi.org/10.1103/PhysRevB.96.024104.

Wen, Kechun, Weiqiang Lv, and Weidong He. 2015. "Interfacial lattice-strain effects on improving the overall performance of micro-solid oxide fuel cells." *Journal of Materials Chemistry A* 3:20031–20050. https://doi.org/10.1039/c5ta03009a.

Willhelm, Daniel, Nathan Wilson, Raymundo Arroyave, Xiaoning Qian, Tahir Cagin, Ruth Pachter, and Xiaofeng Qian. 2022. "Predicting Van der Waals Heterostructures by a Combined Machine Learning and Density Functional Theory Approach." *ACS Applied Materials & Interfaces* 14, no. 22 (June): 25907–25919. https://doi.org/10.1021/acsami.2c04403.

Wolf, D., and S. Yip. 1993. *Materials Interfaces: Atomic-Level Structure and Properties.* Netherlands: Springer.

Wu, Tiantian, Tejs Vegge, and Heine Anton Hansen. 2019. "Improved Electrocatalytic Water Splitting Reaction on CeO_2 (111) by Strain Engineering." *ACS Catalysis* 9:4853–4861.

Xu, Mingsheng, Tao Liang, Minmin Shi, and Hongzheng Chen. 2013. "Graphene-Like Two-Dimensional Materials." *Chemical Reviews* 113:3766–3798. https://doi.org/10.1021/cr300263a.

Xu, Shuigang, Yu Han, Xiaolong Chen, Zefei Wu, Lin Wang, Tianyi Han, Weiguang Ye, et al. 2015. "Van der Waals Epitaxial Growth of Atomically Thin Bi_2Se_3 and Thickness-Dependent Topological Phase Transition." *Nano Letters* 15 (4): 2645–2651. https://doi.org/10.1021/acs.nanolett.5b00247.

Xu, Xin, Yuzi Liu, Jie Wang, Dieter Isheim, Vinayak P. Dravid, Charudatta Phatak, and Sossina M. Haile. 2020. "Variability and origins of grain boundary electric potential detected by electron holography and atom-probe tomography." *Nature Materials* 19 (8): 887–893. https://doi.org/10.1038/s41563-020-0656-1.

Y., Lei Y., Y. Ito, N. D. Browning, and T. J. Mazanec. 2002. "Segregation effects at grain boundaries in fluorite-structured ceramics." *Journal of the American Ceramic Society* 85 (9).

Yamaguchi, Masafumi, Tatsuya Takamoto, Kenji Araki, and Nicholas Ekins-Daukes. 2005. "Multi-junction III–V solar cells: current status and future potential." *Solar Energy* 79 (1): 78–85. https://doi.org/10.1016/j.solener.2004.09.018.

Yang, Shengfeng, and Youping Chen. 2015. "Concurrent atomistic and continuum simulation of bi-crystal strontium titanate with tilt grain boundary." *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Sciences* 471 (2175): 20140758. https://doi.org/10.1098/rspa.2014.0758.

Ye, F., C. Y. Yin, D. R. Ou, and T. Mori. 2014. "Relationship between lattice mismatch and ionic conduction of grain boundary in YSZ." *Progress in Natural Science-Materials International* 24 (1): 83–86.

Yu, G. L., R. V. Gorbachev, J. S. Tu, A. V. Kretinin, Y. Cao, R. Jalil, F. Withers, et al. 2014. "Hierarchy of Hofstadter states and replica quantum Hall ferromagnetism in graphene superlattices." *Nature Physics* 10 (7): 525–529. https://doi.org/10.1038/nphys2979.

Yu, Yifei, Chun Li, Yi Liu, Liqin Su, Yong Zhang, and Linyou Cao. 2013. "Controlled scalable synthesis of uniform, high-quality monolayer and few-layer MoS_2 films." *Scientific Reports* 3. https://doi.org/10.1038/srep01866.

Yuan, Fenglin, Bin Liu, Yanwen Zhang, and William J. Weber. 2016. "Segregation and Migration of the Oxygen Vacancies in the Σ 3 (111) Tilt Grain Boundaries of Ceria." *Journal of Physical Chemistry C* 120 (12): 6625–6632. https://doi.org/10.1021/acs.jpcc.6b00325.

Yuan, Fenglin, Yanwen Zhang, and William J. Weber. 2015. "Vacancy-vacancy interaction induced oxygen diffusivity enhancement in undoped nonstoichiometric ceria." *Journal of Physical Chemistry C* 119:13153–13159.

Yun, Won Seok, S. W. Han, Soon Cheol Hong, In Gee Kim, and J. D. Lee. 2012. "Thickness and strain effects on electronic structures of transition metal dichalcogenides: 2H-MX_2 semiconductors (M = Mo, W; X = S, Se, Te." *Physical Review B* 85 (033305). https://doi.org/10.1103/PhysRevB.85.033305.

Zeng, Hualing, Gui Bin Liu, Junfeng Dai, Yajun Yan, Bairen Zhu, Ruicong He, Lu Xie, et al. 2013. "Optical signature of symmetry variations and spin-valley coupling in atomically thin tungsten dichalcogenides." *Scientific Reports* 3 (30): 11. https://doi.org/10.1038/srep01608.

Zhang, Guanhua, Jianmei Yuan, Yuliang Mao, and Yunqing Huang. 2021. "Two-dimensional Janus material $MoS_{2(1-x)}Se_{2x}$ ($0 < x < 1$) for photovoltaic applications: A machine learning and density functional study." *Computational Materials Science* 186. https://doi.org/10.1016/j.commatsci.2020.109998.

Zhang, L., N. Xu, X. Li, S. Wang, K. Huang, W. H. Harris, and W. K. S. Chiu. 2012. "High CO_2 permeation flux enabled by highly interconnected three-dimensional ionic channels in selective CO_2 separation membranes." *Energy & Environmental Science* 5 (8): 8310–8317.

Zhang, Lei, Zhenjingfeng Yang, Tian Gong, Ruikun Pan, Huide Wang, Zhinan Guo, Han Zhang, and Xiao Fu. 2020. "Recent advances in emerging Janus two-dimensional materials: From fundamental physics to device applications." *Journal of Materials Chemistry A* 8 (18): 8813–8830. https://doi.org/10.1039/d0ta01999b.

Zhang, Leining, Jichen Dong, and Feng Ding. 2021. "Strategies, Status, and Challenges in Wafer Scale Single Crystalline Two-Dimensional Materials Synthesis." *Chemical Reviews* 121 (11): 6321–6372. https://doi.org/10.1021/acs.chemrev.0c01191.

Zhang, Rui, and Rebecca Cheung. 2016. "Two-dimensional Materials - Synthesis, Characterization and Potential Applications: Mechanical Properties and Applications of Two-Dimensional Materials." Chap. 10, edited by Pramoda Kumar Nayak. London: IntechOpen. https://doi.org/10.5772/64017.

Zhang, Y., and W. Yang. 1998. *Physical Review Letters* 80.

Zhang, Yunzhen, Han Ye, Zhongyuan Yu, Yumin Liu, and Yinfeng Li. 2019. "First-principles study of square phase MX_2 and Janus MXY (M=Mo, W; X, Y=S, Se, Te) transition metal dichalcogenide monolayers under biaxial strain." *Physica E: Low-Dimensional Systems and Nanostructures* 110:134–139. https://doi.org/10.1016/j.physe.2019.02.009.

Zhou, Jun, Lei Shen, Miguel Dias Costa, Kristin A Persson, Shyue Ping Ong, Patrick Huck, Yunhao Lu, et al. 2019. "2DMatPedia, an open computational database of two-dimensional materials from top-down and bottom-up approaches." *Scientific data* 6 (1): 86. https://doi.org/10.1038/s41597-019-0097-3.

Zhu, Houyu, Hou Yongchun, Ren Hao, Liu Dongyuan, Li Xin, Zhao Lianming, Chi Yuhua, and Guo Wenye. 2020. "Theoretical investigation of H2 oxidation mechanisms over pristine and Sm-doped CeO_2 (111) surfaces." *Applied Surface Science* 511:145388.

Zhuang, Houlong L., and Richard G. Hennig. 2013. "Computational Search for Single Layer Transition-Metal Dichalcogenide Photocatalysts." *Journal of Physical Chemistry C* 117 (40): 20440–20445. https://doi.org/10.1021/jp405808a.

Zhuang, Houlong L., Michelle D. Johannes, Arunima K. Singh, and Richard G. Hennig. 2017. "Doping-controlled phase transitions in single-layer MoS_2." *Physical Review B* 96 (16): 1–8. https://doi.org/10.1103/PhysRevB.96.165305.

Zhuang, Houlong L., Arunima K. Singh, and Richard G. Hennig. 2013. "Computational discovery of single-layer III-V materials." *Physical Review B - Condensed Matter and Materials Physics* 87 (16): 1–8. https://doi.org/10.1103/PhysRevB.87.165415.

Zouvelou, N., X. Mantzouris, and P. Nikolopoulos. 2008. "Surface and grain-boundary energies as well as surface mass transport in polycrystalline CeO_2." *Materials Science and Engineering: A* 495 (1-2): 54–59. https://doi.org/10.1016/j.msea.2007.10.105.

Zur, A., and T. C. McGill. 1984. "Lattice match: An application to heteroepitaxy." *Journal of Applied Physics* 55 (2): 378–386. https://doi.org/10.1063/1.333084.

APPENDIX A

SUPPLEMENTARY MATERIAL FOR CHAPTER 3

Article	U (eV)	E_{xc}	E_{gap} (eV)	a_0 (Å)
This work	5	PBE	2.1	5.494
ref Koettgen and Martin (2020)	5	PBE	–	5.49
ref Koettgen and Martin (2019)	5	PBE	–	5.49
ref Zhu et al. (2020)***	5	PBE	2.3	5.438
ref Pratik P. Dholabhai et al. (2010a)	5	PBE	2.06	5.494
ref Wu, Vegge, and Hansen (2019)	4.5	PBE	1.8**	5.497

Table A.1. A table comparing this works parameters with recent literature. ** estimated from a DOS plot. *** low k-point density.

A.1 Grain Boundary Structure and Symmetry

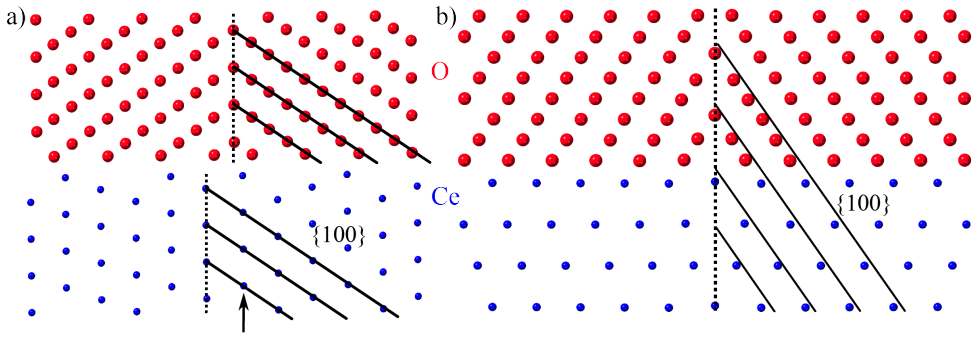

Figure A.1. (a) The Σ3 (111)/[$\bar{1}$01] xz plane for the anion (top) and cation (bottom) sublattices with the {100} planes indicated by the solid black lines. The red spheres are the oxygen ions while the blue spheres are the cerium ions. The dotted black line indicates the {111} mirror plane for each sublattice. Core grain boundary cation sites lie on and to the right of the cation mirror plane. The arrow indications the fully coordinated cation sites (b) The Σ3 (121)/[$\bar{1}$01] xz plane for the anion and cation sublattices with the {100} planes indicated by the solid black lines. The dotted black line indicates the {112} mirror plane. The core grain boundary cation sites are those which lie on either side of the cation mirror plane.

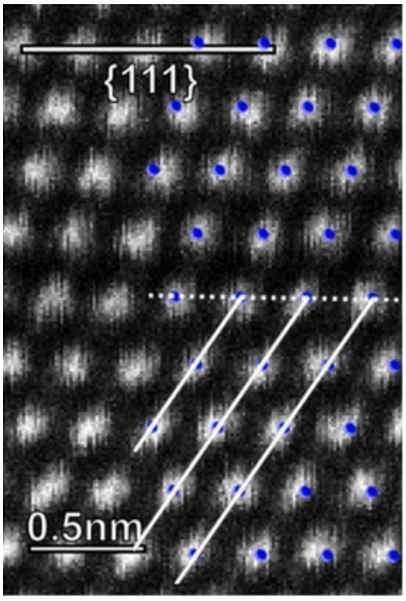

Figure A.2. A high-angle annular dark field scanning transmission electron microscopy image of CeO_2 $\Sigma3$ (111)/[$\bar{1}$01] grain boundary shown in grayscale from B. Feng et al. (2012). The micrograph is superimposed with the CeO_2 $\Sigma3$ (111)/[$\bar{1}$01] grain boundary cation lattice considered in this work where the blue spheres represent the Ce cations. Figure adapted with permission from AIP Publishing.

The $\Sigma3$ (111)/[$\bar{1}$01] grain boundary is a coherent twin boundary possessing separate mirror planes for the cation and anion sublattices indicated by the dashed lines in Appendix Figure A.1a. Each (111) plane has 4 symmetrically equivalent cation sites resulting in 1 crystallographically unique cation site per plane.

The $\Sigma3$ (121)/[$\bar{1}$01] grain boundary is a coherent twin boundary possessing 1 mirror plane for each sublattice indicated by the dashed line in Appendix Figure A.1b. Unlike the $\Sigma3$ (111)/[$\bar{1}$01] grain boundary, the $\Sigma3$ (121)/[$\bar{1}$01] grain boundary has a discontinuity in the cation sublattice seen in Appendix Figure A.1b. This discontinuity results in the more open space structure seen in this grain boundary. Interestingly, the anion sublattice is continuous across all interfaces. The search for unique solute sites may be confined to half the grain boundary due to the presence of a single mirror plane. Each (112) plane in the $\Sigma3$ (121)/[$\bar{1}$01] grain boundary has only 1 unique cation site per plane. Cation sites are defined in Chapter 3. The xy-plane (i.e. the grain boundary plane) is continuous and aligned between both grains for each grain boundary structure model, see insets in Chapter 3 Figure 3.1b and Figure 3.1c.

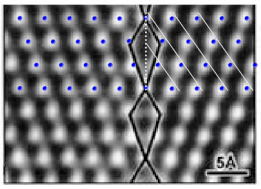

Figure A.3. The grayscale image shows the high resolution transmission electron microscopy image of YSZ $\Sigma 3$ (121)/[$\bar{1}$01] grain boundary. The CeO_2 $\Sigma 3$ (121)/[$\bar{1}$01] grain boundary structure considered in this work is overlaid on the micrograph where blue spheres represent Ce cations. Figure adapted from Shibata et al. (2003) with permission from Taylor & Francis Ltd.

A.2 Atomic Coherency and Solute Configuration Locations

As seen in Chapter 3 Figure 3.1c, the region possessing strained bonds is minimal which is attributed to the atomic coherency of the $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary. This is further illustrated by Chapter 3 Figure 3.1b and Appendix Figure A.1a. Interestingly, the strained cations are fully coordinated and lay along the cation mirror plane. There are 4 coordination-deficient cation sites in the $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary and they are located to the right of the cation mirror plane indicated by the black arrow in Appendix Figure A.1a. Preferential strain of the fully coordinated cations results from the excess volume created from the structural oxygen vacancies that allow the oxygen atoms to relax towards the grain boundary without significantly influencing the bond length of the nearby coordination-deficient Ce-O ions.

In the $\Sigma 3$ (121)/[$\bar{1}$01] grain boundary, the discontinuity of the cation sublattice across the interface locally distorts the lattice introducing significant bond strain seen in Appendix Figure A.1b and Chapter 3 Figure 3.1e. The $\Sigma 3$ (121)/[$\bar{1}$01] grain boundary has 4 coordinated-deficient and 2 highly strained fully-coordinated cation sites further disrupting the bonds at/near the interface. The above mentioned effects explain the significantly strained bonds surround the grain boundary region as compared to the more coherent $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary.

A.3 Stable Bulk AEM Phases

To identify the stable bulk oxides for the AEM metals, the Materials Project (Ong et al. 2013) database is searched for compounds matching the constituent elements from which only the compounds with the correct stoichiometry (MO) are used for comparison. For BeO compounds, there are a total of 9 oxide compounds with

Mat.	mp-id	Sg.	M-O (Å)	M-coord	E (eV/f.u.)
BeO	mp-2542	$P6_3mc$	1.66	tetrahedral	-14.2574
MgO	mp-1265	$Fm\bar{3}m$	2.13	octahedral	-11.9333
CaO	mp-2605	$Fm\bar{3}m$	2.42	octahedral	-12.9008
SrO	mp-2472	$Fm\bar{3}m$	2.6	octahedral	-12.0736
BaO	mp-1342	$Fm\bar{3}m$	2.81	octahedral	-11.8279
CeO_2	mp-20194	$Fm\bar{3}m$	2.38	cubic	-24.373

Table A.2. The table lists the materials composition, Materials Project id's (mp-id), space group (Sg.), M-O bond length in Å where M = Be, Mg, Ca, Sr, and Ba, the coordination of the cation (M-coord) and the energy (E in eV/f.u.) of MO_x used to compute ΔE_{GB}.

the correct stoichiometry. Of these 9 structures, the structures considered here are reported to have been experimentally synthesized. There are only 3 such BeO crystal structures. Their mp-id's and space groups are: 2542, 7599, 1794 and $P6_3mc$, $P4_2/mnm$, $Fm\bar{3}m$, respectively. The structures are listed from the lowest $E_{above\ hull}$ to the largest. $E_{above\ hull}$ is the energy of decomposition of a material into the set of most stable materials at this chemical composition. The structure corresponding to mp-id 2542 is the most energetically stable structure. The Be coordination is 4 fold with three bond lengths at 1.65 Å and one bond length 1.66 Å. The coordination polyhedra for Be is tetrahedral with BeO_4.

For MgO compounds, there are a total of 23 oxide structures. Of these 23 structures, 3 structures are experimentally derived MgO structures. Their mp-id's and space groups are: 1265, 549706, 1009129 and $Fm\bar{3}m$, $P6_3mc$, $P\bar{6}m2$, respectively. The most energetically stable MgO structure is Halite with mp-id 1265 which is the rock salt crystal structure. The coordination of Mg atoms is octahedral with 6 equivalent bonds of length 2.13 Å. However, as discussed in Chapter 3, at the ceria grain boundary a 4-fold coordination is observed for the Mg ion. The second most stable MgO compound in the MP database is a wurtzite structured oxide with mp-id 549706 which possesses 4-fold coordinated Mg with bond lengths of 1.99 Å and 2.02 Å. The Mg coordination polyhedra is tetrahedral with MgO_4.

For CaO compounds, there are a total of 7 oxide structures. Of these 7 structures, there are 3 structures which are experimentally derived CaO crystal structures which have mp-id's 2605, 1079707, and 1064492. These candidate structures have space-groups $Fm\bar{3}m$, Cmce, and I4/mmm, respectively. The most stable CaO compound on the MP database is a Halite/Rock Salt structure with Ca atoms coordinated with 6 equivalent bonds of length 2.42 Å. The Ca coordination polyhedra is octahedral CaO_6.

For SrO compounds, there are a total of 4 oxide structures. Of these 4 structures,

only 1 structure is an experimentally derived SrO crystal structure with mp-id 2472 and a corresponding space-group $Fm\bar{3}m$. This is a Halite/Rock Salt structure with Sr having 6-fold coordinated and bond lengths of 2.6 Å. The coordination polyhedra for Sr is octahedral, SrO_6.

For BaO compounds, there are a total of 11 oxide structures. Of these 11 structures, there are 3 structures which are experimentally derived BaO crystal structures which have mp-id's 1342, 1008500 (MAX-Phase), 7487 (2 dimensional material). Due to the low dimensionality of the latter 2 crystal structures only the first crystal structure is a viable structure which has a space group $Fm\bar{3}m$ which also happens to be the most stable form. The structure is Halite/Rock Salt with the Ba coordinated with 6 equivalent bonds of length 2.81 Å. The coordination polyhedra is octahedral with BaO_6.

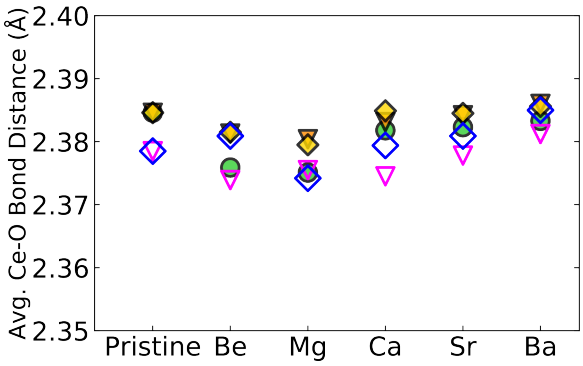

Figure A.4. The supercell averaged Ce-O bond distances for each grain boundary-solute structure model. The $\Sigma 3$ (111)/[$\bar{1}$01] ($\Sigma 3$ (121)/[$\bar{1}$01]) grain boundary solute sites are indicated with open (filled) markers. Chapter 3 Figure 3.2 further illustrates these markers for all the sites in each grain boundary. The general trend is that the average Ce-O bond distance is a minimum for the smaller solutes and increases with increasing solute radii.

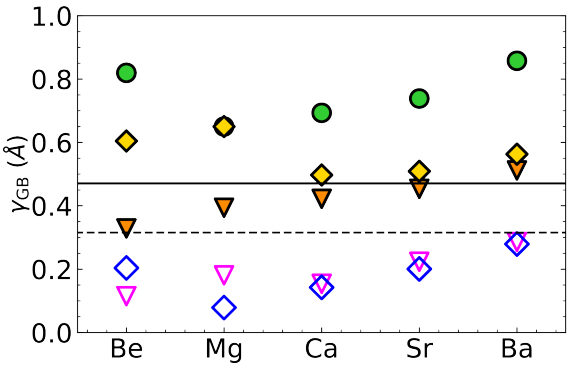

Figure A.5. The z-axis grain boundary expansion (γ_{GB}) for each grain boundary-solute structure model. Dashed (solid) line corresponds to the expansion for the $\Sigma 3$ $(111)/[\bar{1}01]$ ($\Sigma 3$ $(121)/[\bar{1}01]$) grain boundary. The $\Sigma 3$ $(111)/[\bar{1}01]$ ($\Sigma 3$ $(121)/[\bar{1}01]$) grain boundary solute sites are indicated with open (filled) markers. The γ_{GB} is highly site and solute dependent. Generally, there is a linear increase in the grain boundary expansion with increasing solute size. However, this trend is not well established for smaller solute sizes due to the propensity for the solute ions to occupy interstitial solute sites resulting in an increased grain boundary expansion.

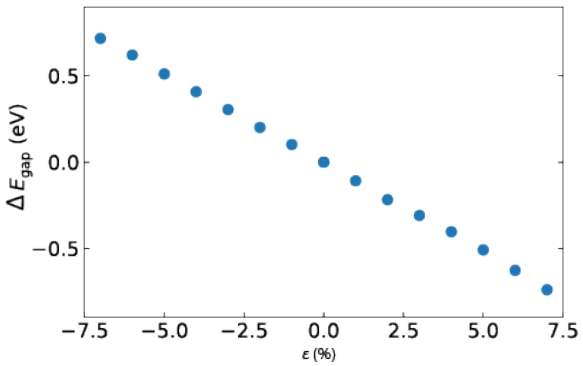

Figure A.6. The change in band gap of CeO_2, ΔE_{gap}, as a function of uniform hydrostatic strain, ϵ, applied on the lattice.

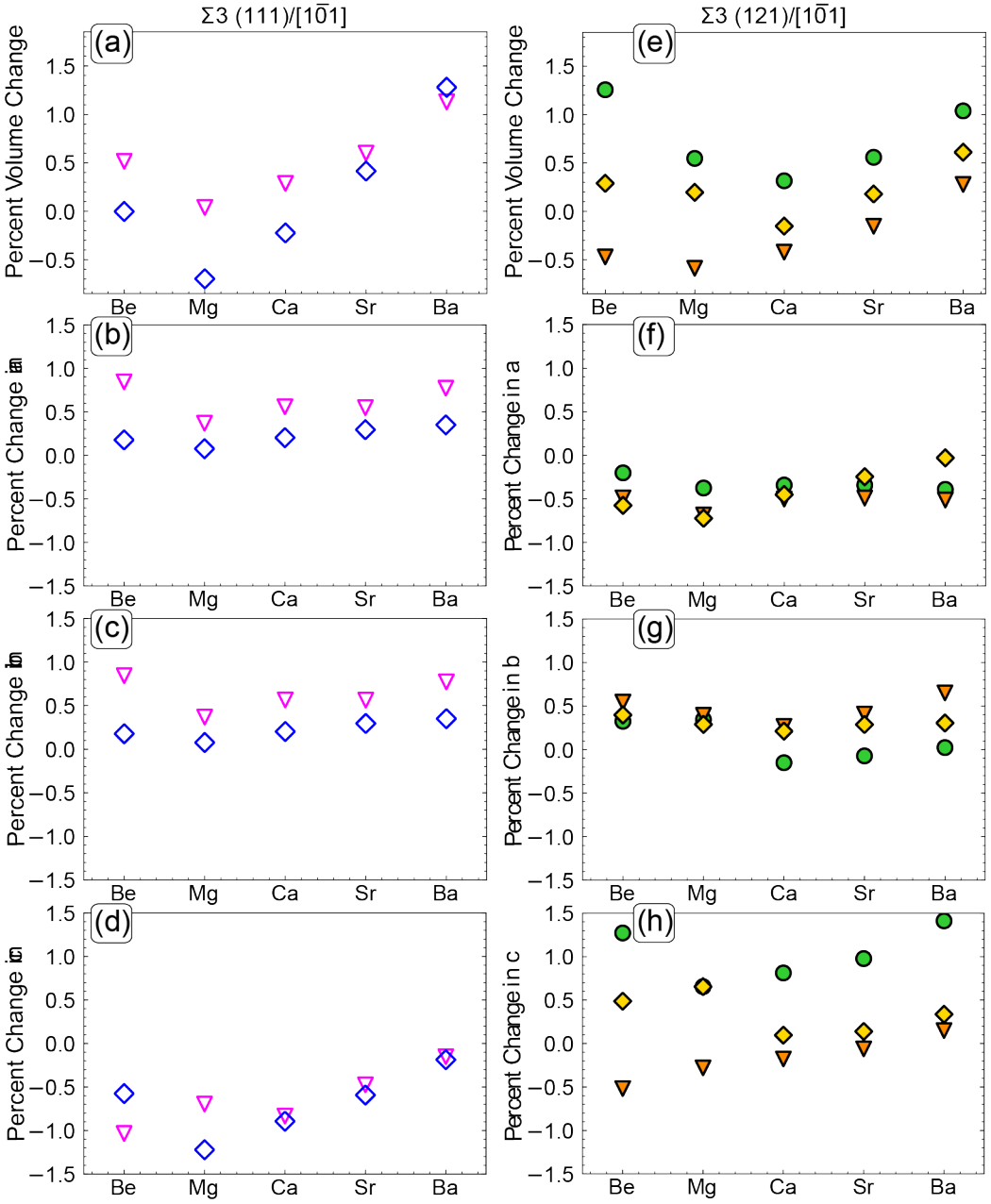

Figure A.7. The percent change in the volume (a, e), **a** (b, f), **b** (c, g), and **c** (d, h) lattice vectors for the Σ3 (111)/[$\bar{1}$01] and Σ3 (121)/[$\bar{1}$01] grain boundary for each grain boundary-solute structure model. Open (solid) markers corresponds to the Σ3 (111)/[$\bar{1}$01] (Σ3 (121)/[$\bar{1}$01]) grain boundary. Percent change is referenced from the respective undoped grain boundary parameters.

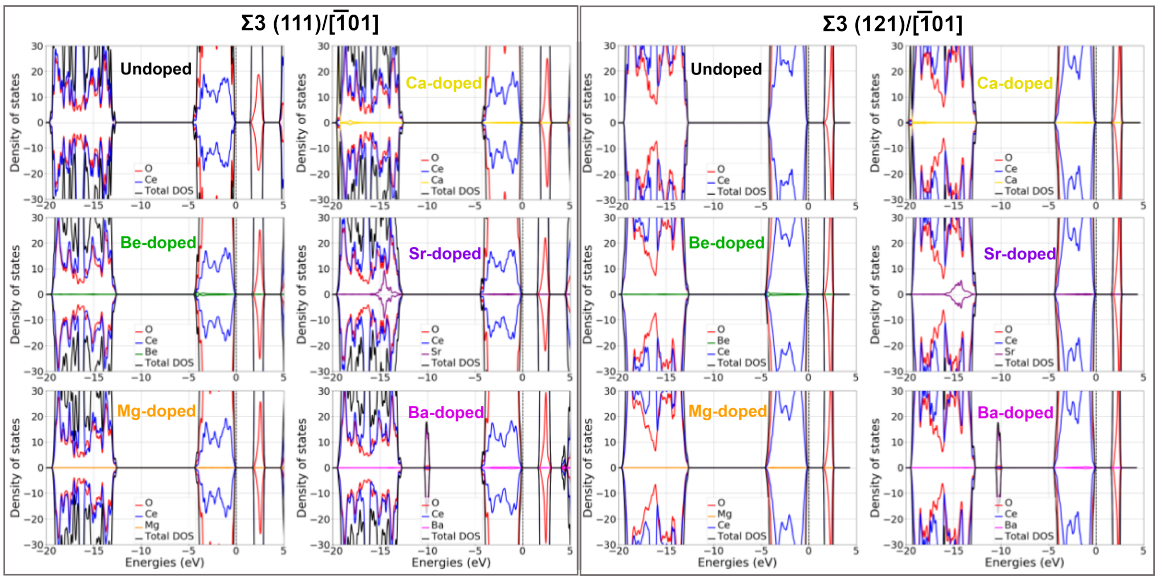

Figure A.8. Element-projected DOS for the lowest energy grain boundary configurations. First two columns are the Σ3 (111)/[$\bar{1}$01] grain boundary DOS and the last 2 columns are the Σ3 (121)/[$\bar{1}$01] grain boundary DOS.

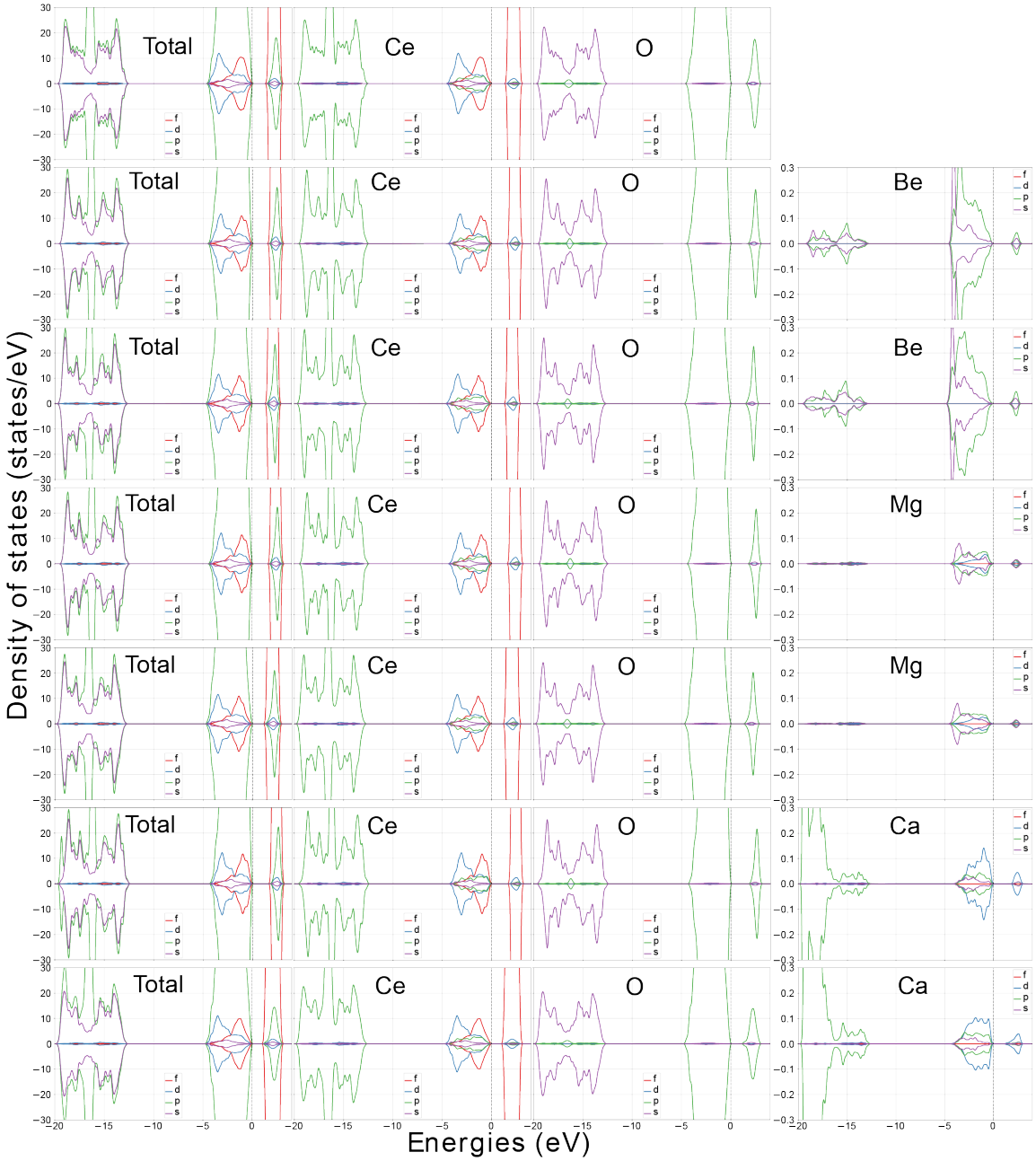

Figure A.9. The orbital-projected DOS for the Σ3 (111)/[$\bar{1}$01] grain boundary. Each row represents a unique grain boundary-solute configuration and each column (from left to right) corresponds to the supercell, Ce, O, and solute lm-projected DOS. Each row (from top to bottom) corresponds to the undoped, Be, Mg, and Ca configurations where the solute configurations are display with the magenta triangle site first, then the blue diamond site.

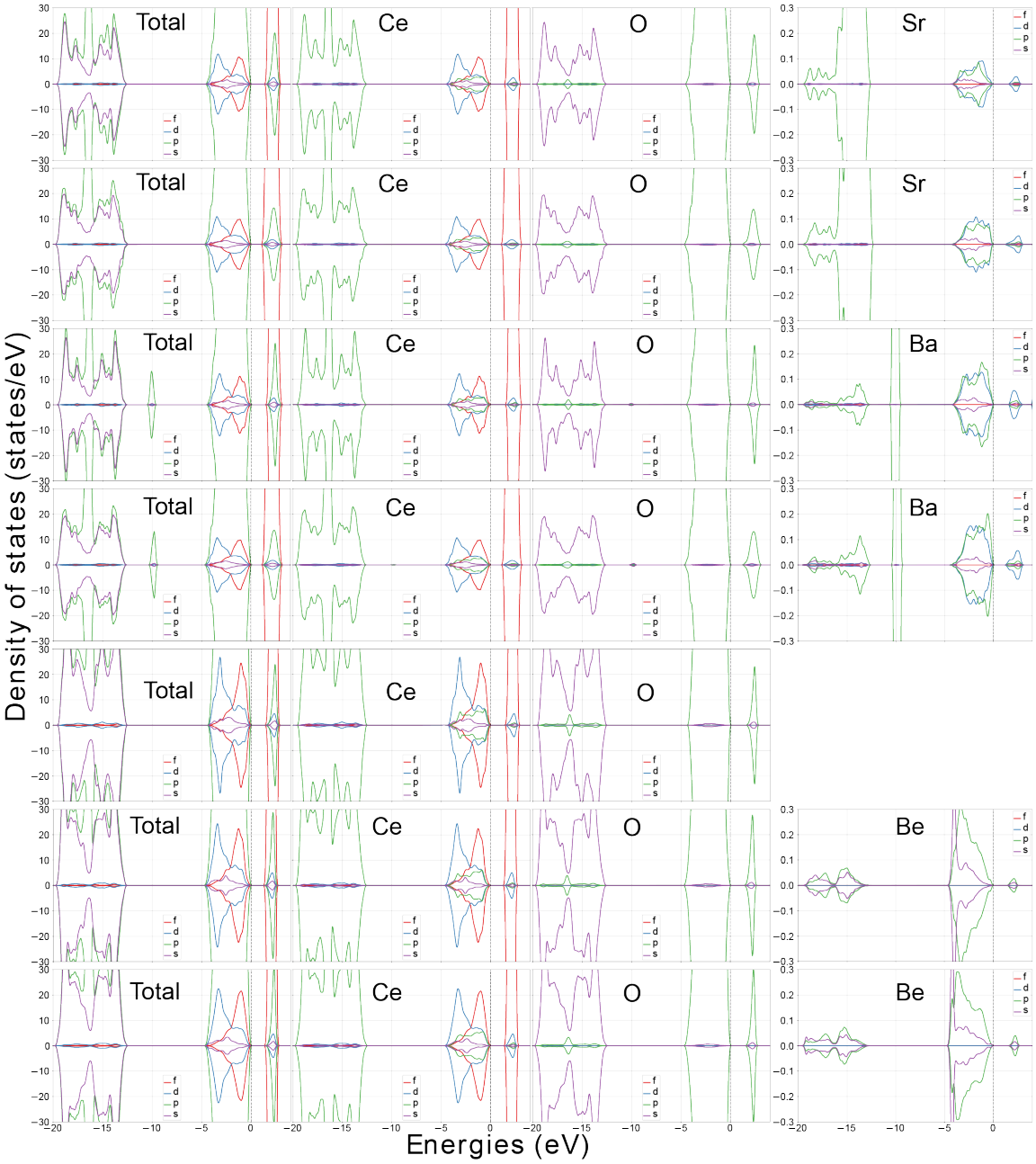

Figure A.10. The orbital-projected DOS for the Σ3 (111)/[$\bar{1}$01] and Σ3 (121)/[$\bar{1}$01] grain boundary. The first 4 rows correspond to the Sr- and Ba- doped Σ3 (111)/[$\bar{1}$01] grain boundaries with the configurations ordered the same as in Appendix Figure A.9 as are all columns listed identically. The following 3 rows corresponds to the Σ3 (121)/[$\bar{1}$01] undoped and first 2 Be configurations, where the solute configurations for the Σ3 (121)/[$\bar{1}$01] grain boundary are displayed with the green circle site first, then the orange triangle site.

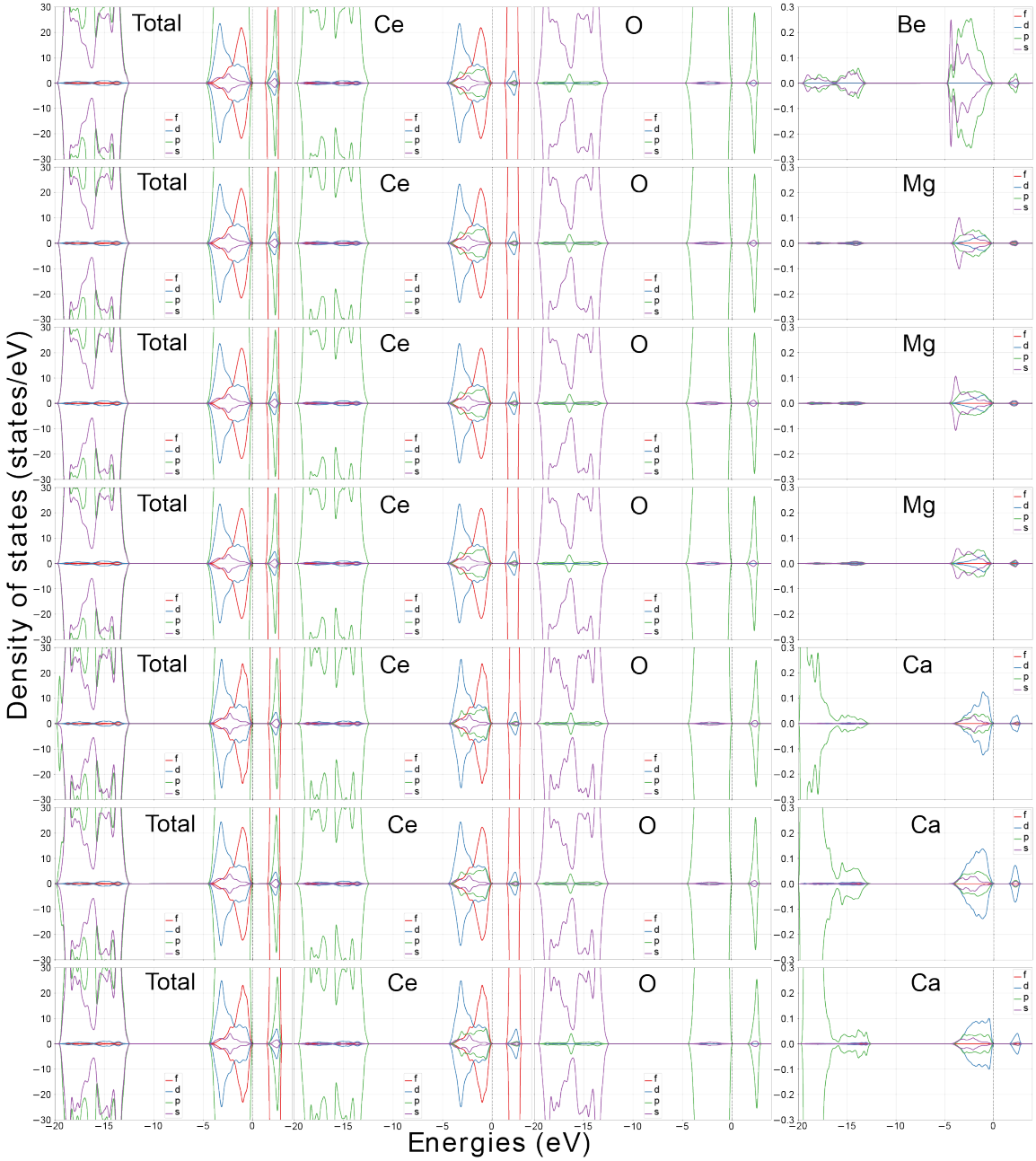

Figure A.11. The orbital-projected DOS for the Σ3 (121)/[$\bar{1}$01] grain boundary. The first row correspond to the yellow diamond Be site, the next two sets of 3 rows correspond to Mg and Ca configurations ordered with the green circle, orange triangle, and yellow diamond site.

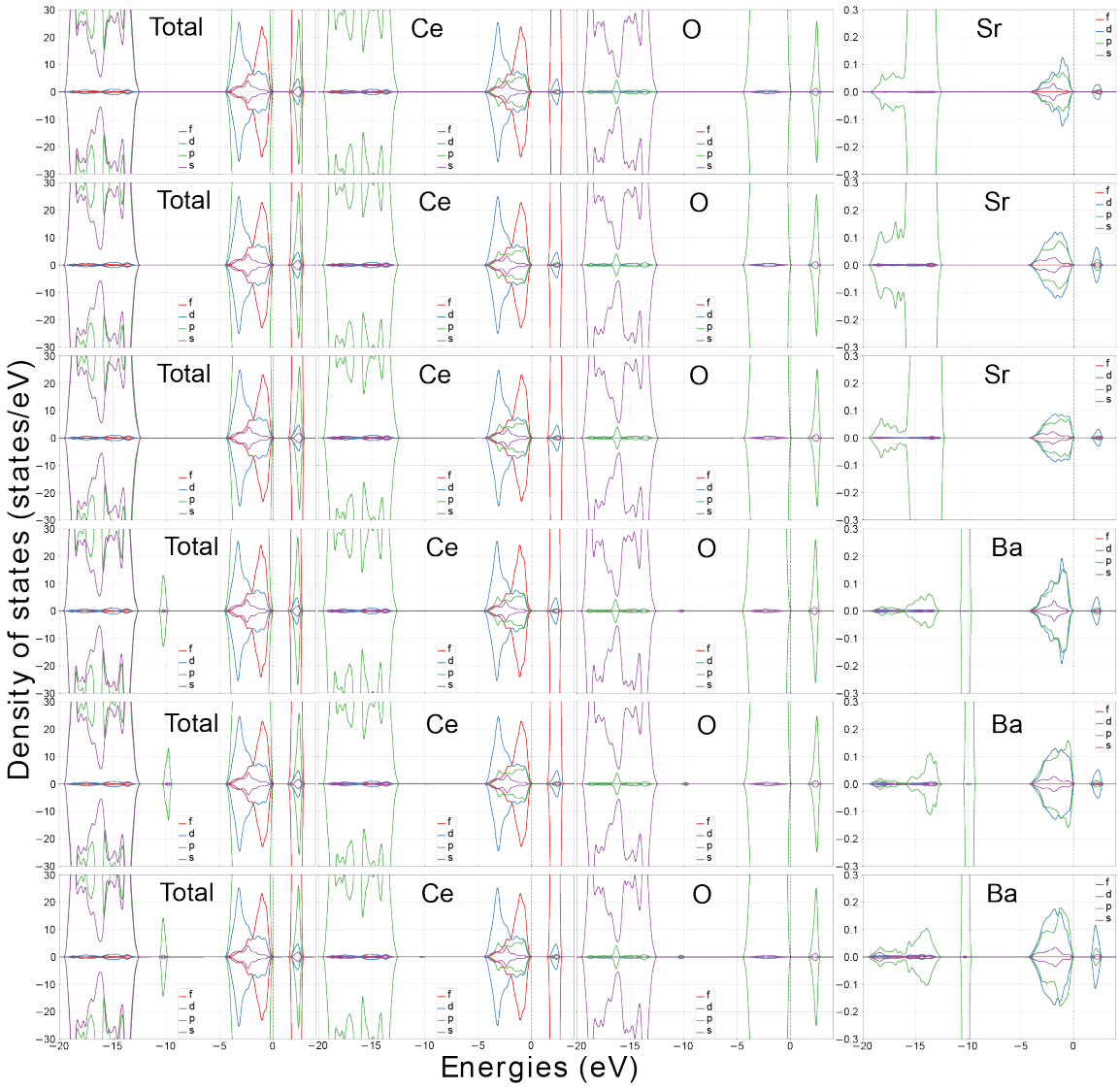

Figure A.12. The orbital-projected DOS for the Σ3 (121)/[$\bar{1}$01] grain boundary. The first 3 rows correspond to the Sr sites while the last 3 rows correspond to the Ba sites where the configurations are listed in the same order described in Appendix Figure A.11.

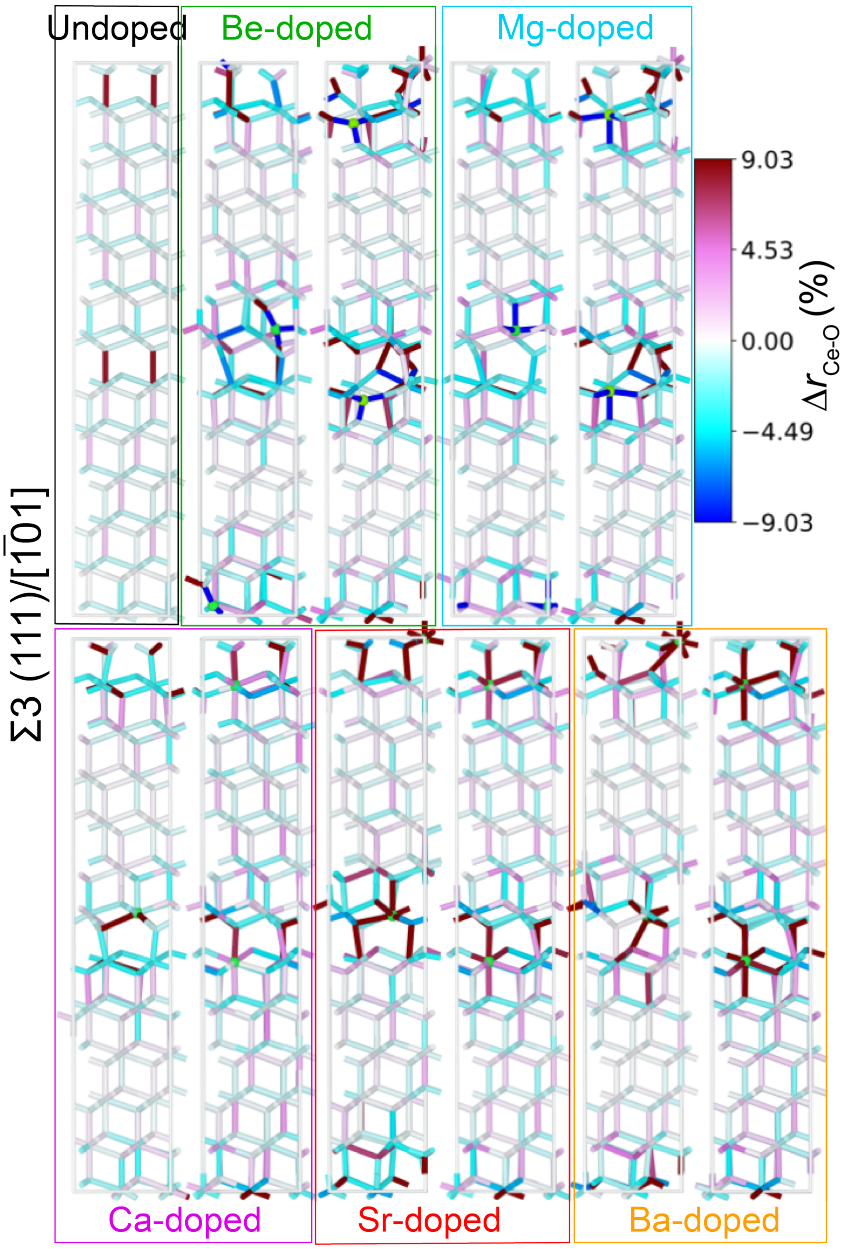

Figure A.13. Percent bond deviation strain maps for the Σ3 (111)/[$\bar{1}$01] grain boundary structure models. The first row contains the undoped, Be, and Mg structures. The configurations are listed in increasing numerical order from left to right. The second row contains Ca, Sr, and Ba doped grain boundary structure models. Color map indicates percent bond deviation referenced using the average Ce-O bond distance between the two grain boundary models. Negative value corresponds to compressed bonds while red corresponds to tensile bonds. Green spheres indicate the solute location.

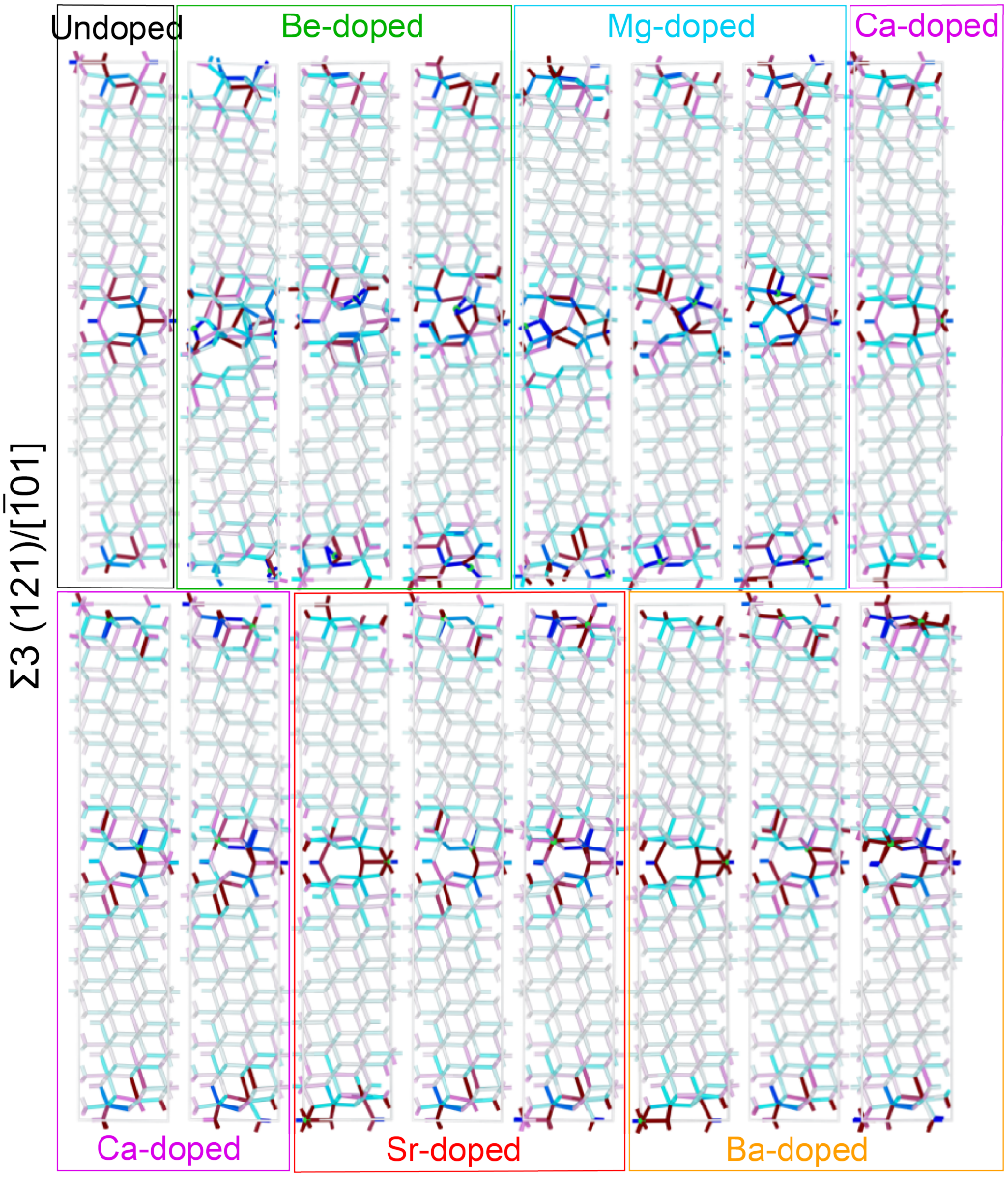

Figure A.14. The percent bond deviation strain maps for the Σ3 (121)/[$\bar{1}$01] grain boundary structure models. The first row contains the undoped, Be, and Mg structures where the configurations are listed in numerical order from left to right with the last model being the first Ca configuration. The second row contains the Ca, Sr, and Ba grain boundary structure models. Color map is the same as in Appendix Figure A.13.

A.4 Electronic Structure of AEM Doped Σ3 (111)/[$\bar{1}$01] Grain Boundaries

Appendix Figure A.15 shows the full DOS for each AEM doped Σ3 (111)/[$\bar{1}$01] grain boundary where the DOS for each configuration is aligned to the low level O *s* states. The solute configuration profiles are plotted together (configuration 0 in red and configuration 1 in blue) to illustrate how the solute's location impacts the DOS. Appendix Figure A.16 shows a close up of the DOS for each grain boundary to discern in which band the DOS shifts.

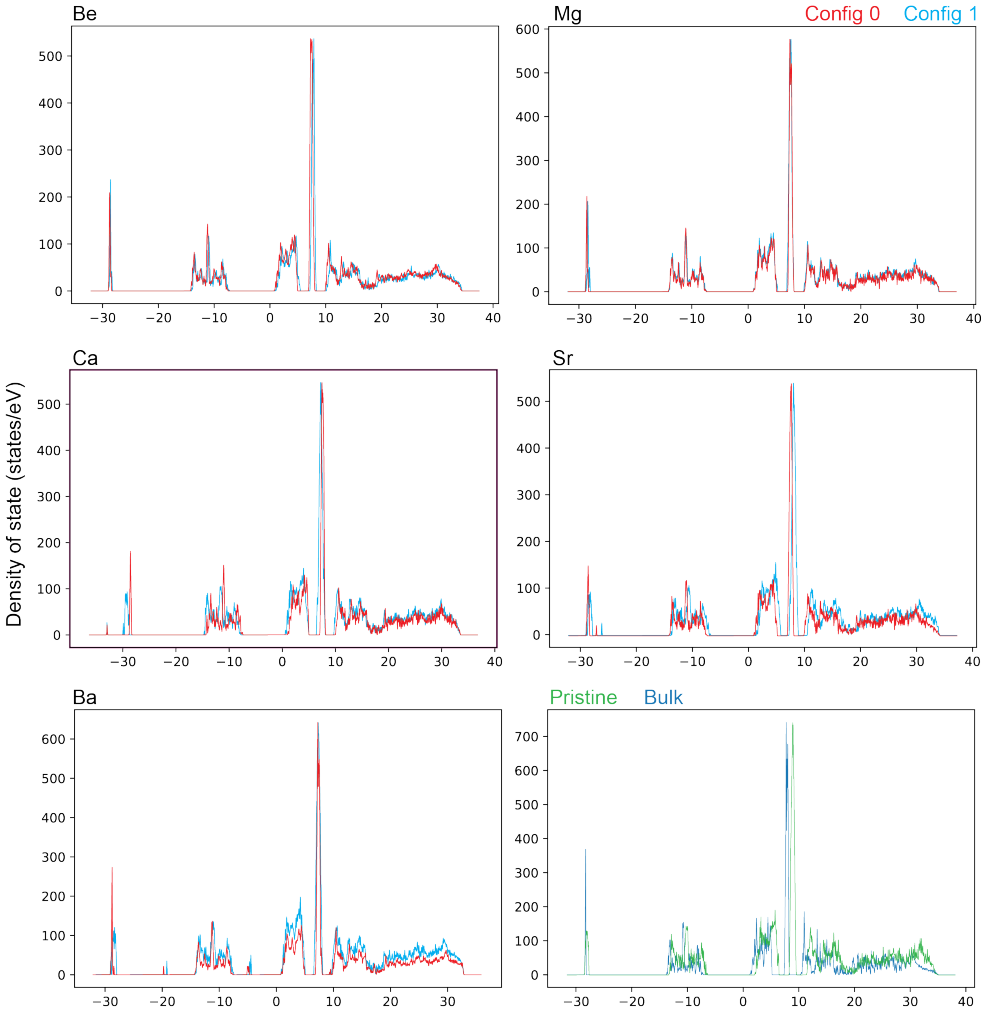

Figure A.15. DOS for AEM doped Σ3 (111)/[$\bar{1}$01] grain boundary aligned to the low lying oxygen *s* states.

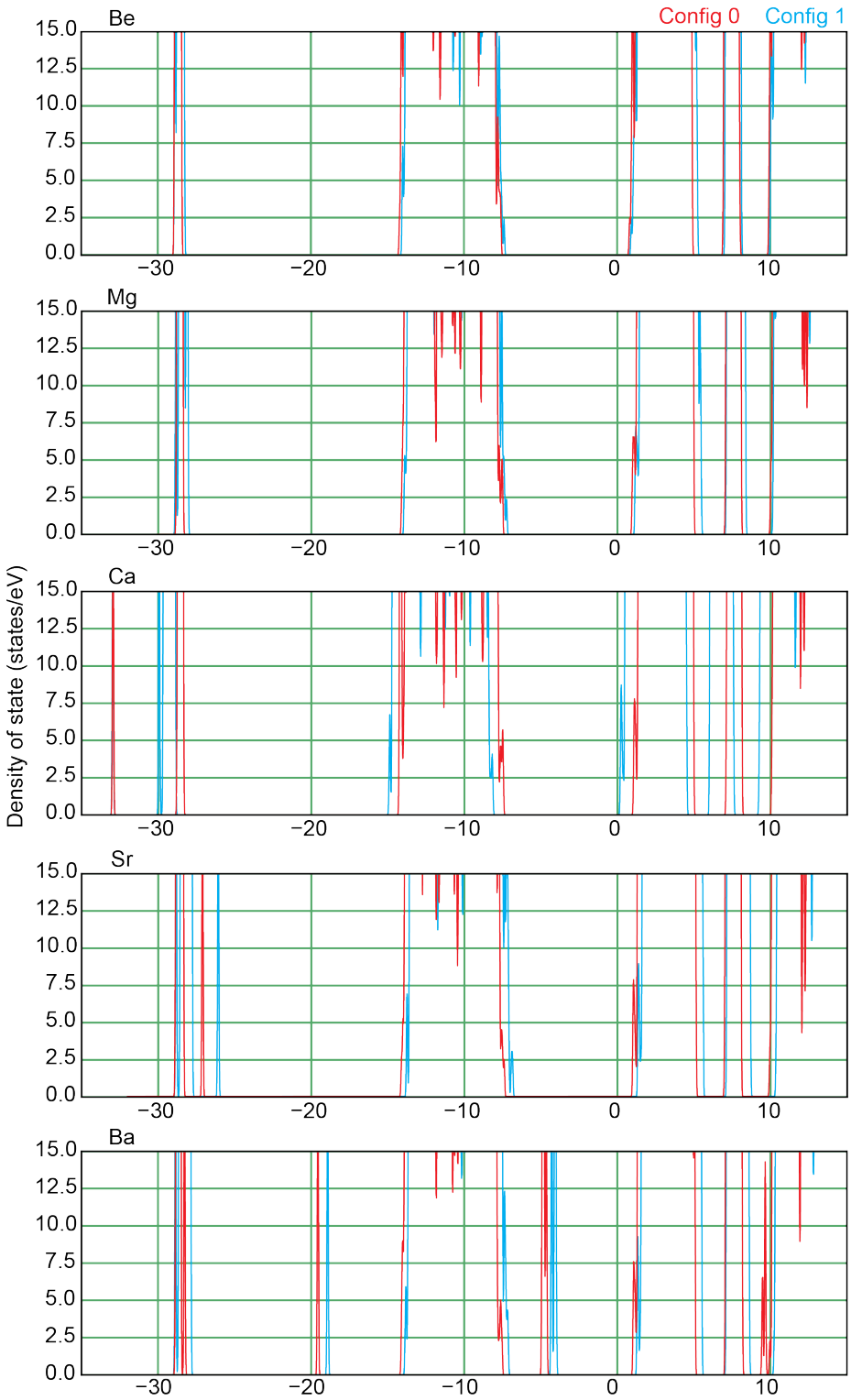

Figure A.16. DOS for AEM doped Σ3 (111)/[$\bar{1}$01] grain boundary aligned to the low lying oxygen *s* states re-scaled to illustrate the changes due to the solute ions.

APPENDIX B

SUPPLEMENTARY MATERIAL FOR CHAPTER 4

B.1 Electronic Structure and Charge Density of O_{vac} Migration in $\Sigma 3$ (111)/[$\bar{1}$01] Grain Boundary

Appendix Figure B.1-B.2 represents the structure model for the undoped $\Sigma 3$ (111)/[$\bar{1}$01] grain boundary overlaid with partial charge density for the bands containing the two $4f$ electrons localized on Ce resulting from the removal of an oxygen ion. The plots below are the corresponding DOS and band structure for each O_{vac} location. The white box represents the oxygen vacancy location while the orange isosurfaces represent the band projected partial charge density for the two localized defect states which can be seen both in the DOS and band structure in the band gap.

One interesting trend to notice in the electronic structure/charge density is that the two electrons do not localize only on the Ce ions near the vacancy. One electron follows the oxygen vacancy while the other localizes on a Ce ion in the grain boundary core. Only when the oxygen vacancy is sufficiently far from the grain boundary core does the second electron localize on a Ce ion near the oxygen vacancy. However, it is unclear if this is due to the size of the supercell and more testing–by explicitly localizing the electron on Ce ions near the oxygen vacancy for each location–is needed to determine the origin of the electron localizing at the grain boundary. However, this finding would not be all to surprising as similar systems (Frechero et al. 2015) have shown electrons preferentially localizing to the grain boundary as well.

One particular challenge in calculating the properties of defective ceria is dealing with the localized electrons. The localized electrons can cause convergence issues and/or localize on Ce ions which may not be the optimal site. In order to overcome these challenges, it is often necessary to obtain pre-converged WAVECAR before relaxing a structure. Additionally, when generating the pre-converged WAVECAR, identifying which Ce ions the electrons localize on and specifying this using the MAGMOM tag to enforce this condition could prevent issues with later simulations. For the calculations involving defective ceria, it was sufficient to simply start the relaxation from a pre-converged WAVECAR. However, issues became apparently in later simulations originating from the localized electron while attempting to use NEB. The method used to over manage that issue is discussed in the relevant section below.

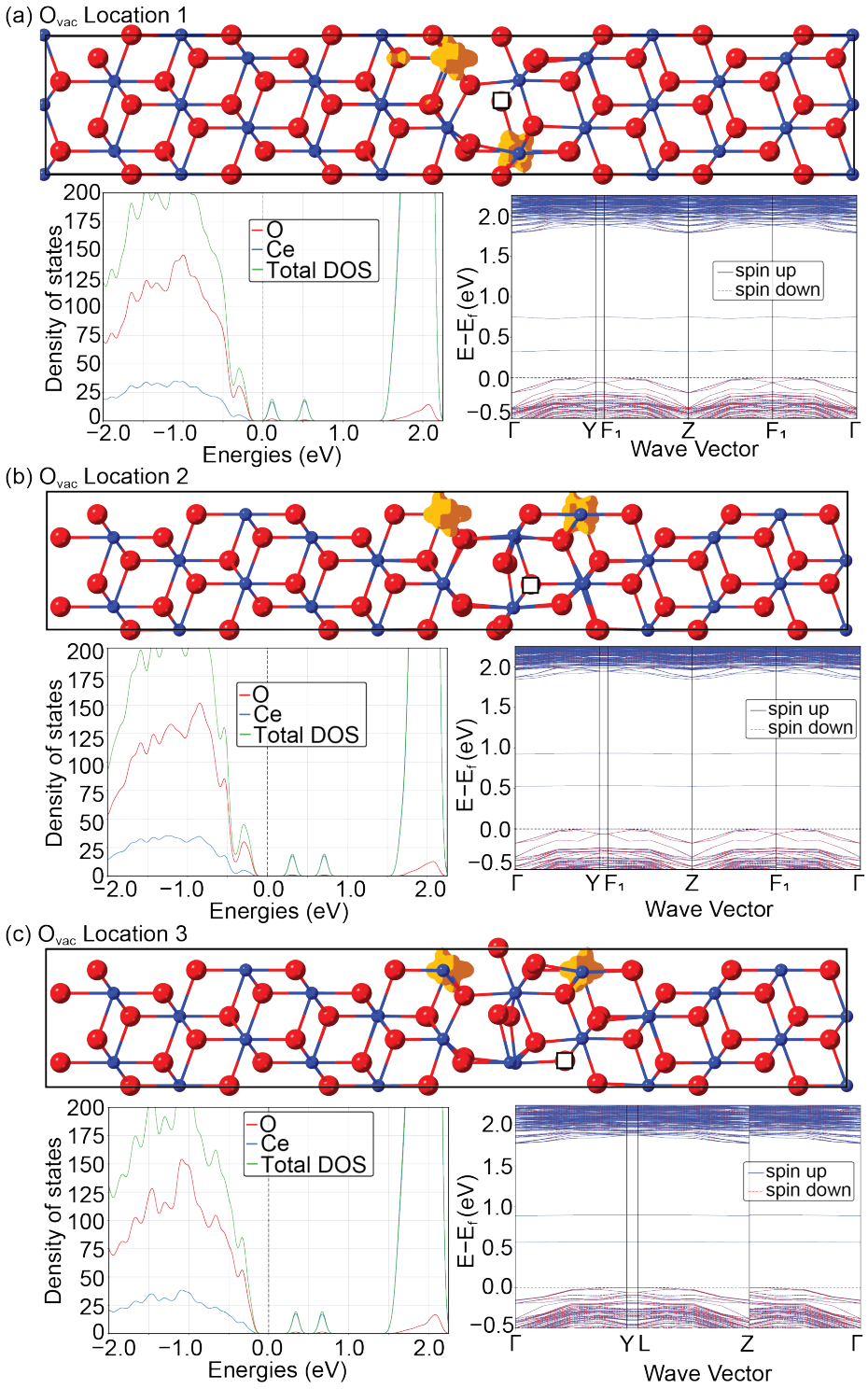

Figure B.1. Structure model overlaid with partial charge density and the corresponding DOS and band structure for each O_{vac} location.

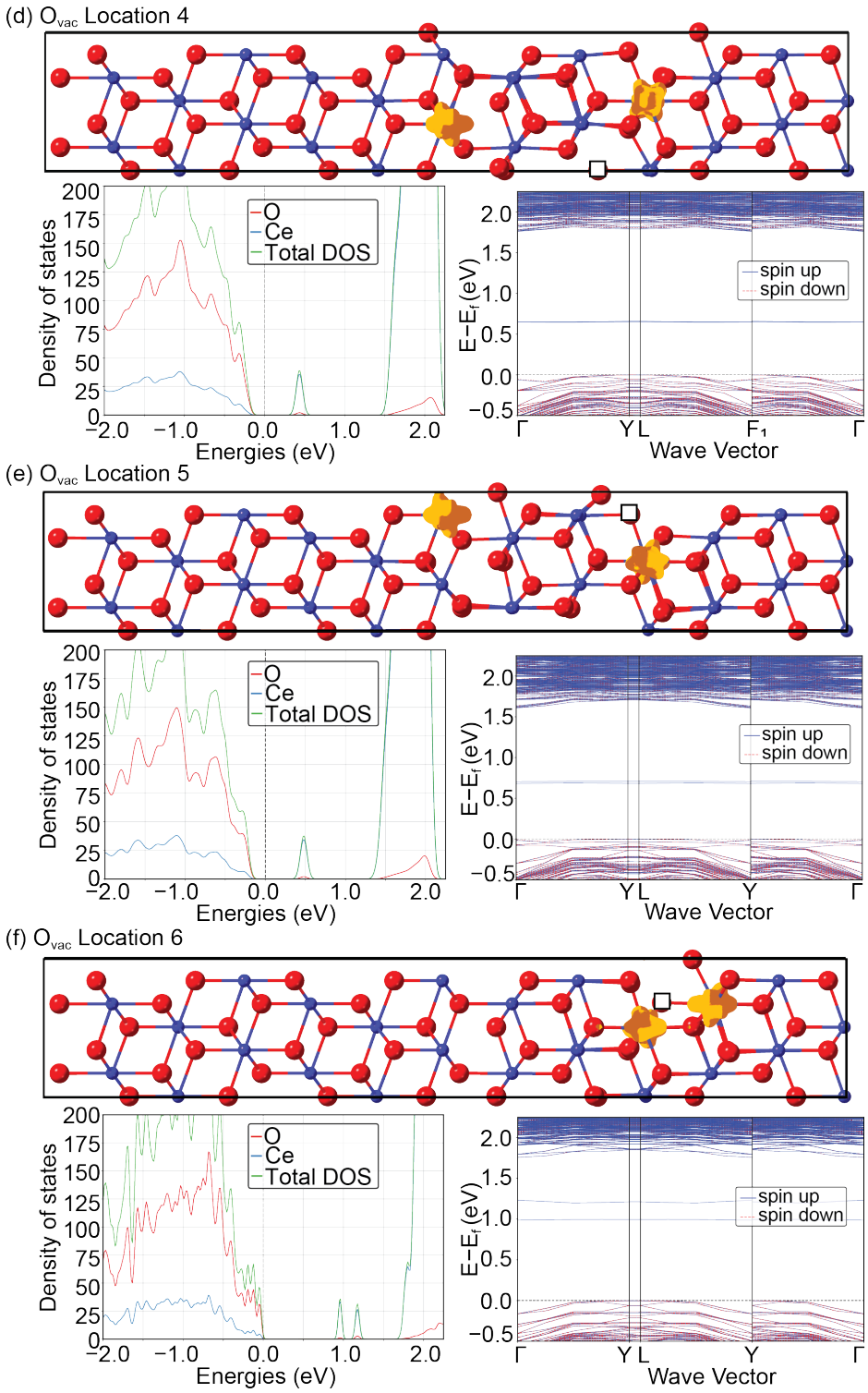

Figure B.2. Structure model overlaid with partial charge density and the corresponding DOS and band structure for each O_{vac} location.

B.2 NEB Methodology for Bulk and Grain Boundary Calculations

Initial testing using the NEB method was done using a 2x2x2 bulk supercell of ceria with one oxygen vacancy. The first NEB path had 3 images and initialized without pre-converging the WAVECAR. These calculations often resulted in highly asymmetric and nonphysical barrier pathways. To ensure the calculation would converging to a more physical solution, a WAVECAR for each image was generated. For the 3 images this solution worked remarkably well. However, increasing the number of images from 3 to 5 generated asymmetric barriers. It is likely originating from the differences between the localized electrons. To solve this prior to generating a pre-converged WAVECAR the 4 Ce ions in which the electrons are localized on must have the magnetic moment specified using the MAGMOM tag to force an more realistic and symmetric arrangement of the electrons during oxygen vacancy migration. Figure B.3 shows the NEB barrier for a bulk 2x2x2 ceria supercell with one oxygen vacancy and 3 images. Additionally, the projected charge density in each image also reflects this symmetry.

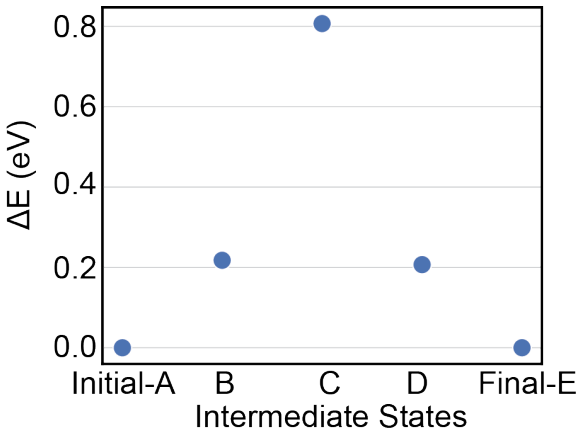

Figure B.3. NEB barrier for a bulk 2x2x2 ceria supercell with one oxygen vacancy and 3 images. Note the barrier is symmetric.

While performing the NEB calculations for the grain boundary structure all calculations would fail to complete even one ionic step unless a pre-convergenced WAVECAR was generated for each image which specified the MAGMOM tag for the Ce ions with the localized electrons. For some unclear reason, not specifying the MAGMOM tag and generating the WAVECAR was not sufficient and the calculation would repeatedly fail. The only method which was successful was generating a new WAVECAR with the MAGMOM tag on the respective Ce ions which could be found by finding the atom number which possess the two localized f states in the OUTCAR.

Other potential (less extreme/time consuming) solutions which worked to converge the NEB calculations was changing from IBRION 2 to 1, specifying a smaller AMIN value (0.01), or using the WAVECAR from the initial and final images as the initial WAVECAR for the images.

APPENDIX C

SUPPLEMENTARY MATERIAL FOR CHAPTER 5

C.1 Default Computational Parameters Set by CMDLInterfaceSet

The initial structures for the 2D materials, the bulk counterpart of the 2D materials, and the substrates are imported from existing materials databases (Ong et al. 2013; Haastrup et al. 2018) or can be input by the users. The structures obtained from databases have been determined by DFT simulations that employ different parameters and exchange-correlation functionals. For consistency, all imported structures are relaxed using the *CMDLInterfaceSet* routine to automate the creation of the DFT input files to perform the DFT calculations using the vdW-DF-optB88 functional. The vdW-DF-optB88 (Rydberg et al. 2003; Bowler and Michaelides 2011) functional accounts for non-local dispersion interactions. It has successfully represented the weak vdW forces in layered materials, accurately reproducing not only the inter-layer spacings but also the dispersion interactions between 2D materials and substrate surfaces (Singh, Zhuang, and Hennig 2014; Singh and Hennig 2014a; Zhuang, Singh, and Hennig 2013).

CMDLInterfaceSet is based on pymatgen's *VASPInputSet* class used to customize VASP calculations. To generate files for VASP calculations, the new class *CMDLInterfaceSet* has all the functionality of the parent pymatgen class but is tailored to perform structural optimizations of 2D-substrate heterostructures and implement vdW-corrections, on-the-fly dipole corrections for slabs, custom k-point mesh grid density, and automated selective dynamics tags for the interface atoms of the 2D-substrate structures. All DFT calculations are performed using the projector-augmented wave method as implemented in the plane-wave code VASP (Kresse and Joubert 1999; Kresse and Furthmüller 1996b; Kresse and Hafner 1993, 1994; Kresse and Furthmüller 1996a). The vdW interactions between the 2D materials and substrates are modeled using the vdW–DF functional with the optB88 exchange functional (Bowler and Michaelides 2011).

The default settings of the *CMDLInterfaceSet* require all calculations to converge to total force per atom of less than 0.02 eV/Å or better and an energy tolerance of 10^{-4} and 10^{-6} eV for supercells with less or greater than 1000 atoms, respectively. These parameters were sufficient to converge all binding energies in the benchmark simulations to 10^{-3} eV/atom. The convergence tests for the binding energy can be found below as Appendix Figure C.1. The adsorption formation energies are performed using a slab geometry with a minimum vacuum spacing of 18 Å which ensures the interactions between the surfaces are negligible. The combined setting for the energy cutoff, k-point density, and total-force per atom provide reasonable accuracy for the $\Delta E^{f}_{\mathrm{vac}}$, ΔE_{b}, and $\Delta E^{f}_{\mathrm{ads}}$ while maintaining a relatively small computational cost for high-throughput calculations. The combined 2D-substrate heterostructures are created with a lattice mismatch $\leq 5\%$, a surface area ≤ 130 Å^2, and an initial 2D-substrate separation distance of 3 Å – 3.4 Å. These heterostructure parameters were selected to

reduce the computational cost of simulating large supercells and minimize the strain in the 2D material.

C.2 Convergence Tests

In order to confirm that the convergence parameters used in these calculations are sufficient, convergence are performed checks at 3 different energy cut off values (10^{-4}, 10^{-5}, and 10^{-6}) for 3 different structures with an increasing number of atoms in the supercell. Convergence in the binding energy to at least 3 significant figures is observed given the default energy cut off of 10^{-4}. The results are shown in Appendix Figure C.1.

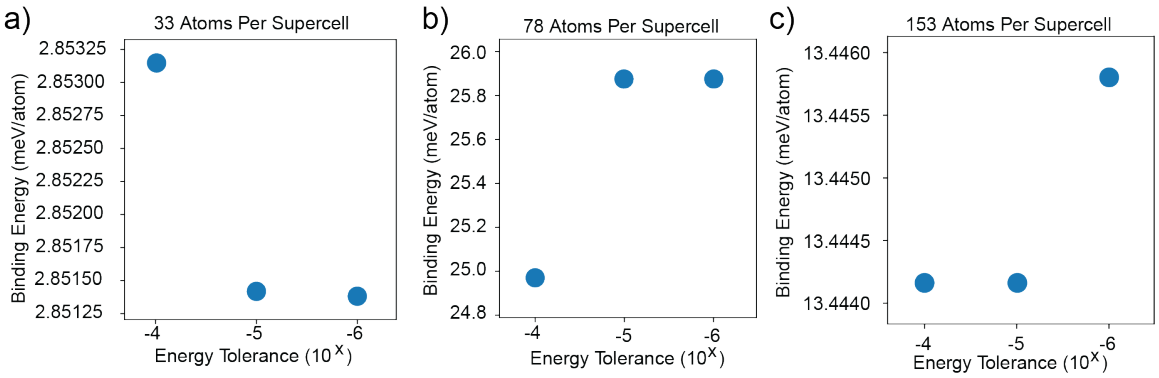

Figure C.1. Binding energy values calculated with 3 different energy tolerance values to test binding energy convergence for supercells with different atom numbers. The supercells contain (a) MoS_2 on (111) Ir with 33 atoms, (b) MoS_2 on (110) Sc containing 78 atoms, and (c) 1*H*-NbO_2 on (111) Ir containing 153 atoms. The binding energy (in eV/atom) is converged to (a) 5 significant figures, (b) 3 significant figures, and (c) 5 significant figures.

C.3 Initialization and Customization of the Workflow

In Chapter 5 Figure 5.4, the *slab_params* dictionary generates a substrate slab with a vacuum spacing of 19 Å and a substrate slab thickness of at least 12 Å. The *h_params* dictionary creates the lattice-matched, symmetry-matched 2D-substrate heterostructures with 3.0 Å z-separation distance between the 2D material and the substrate surface. The *h_params* dictionary also sets the maximum allowed lattice-mismatch along **ab** to be less than 5%, a surface area less than 130 $Å^2$, sets the

selective dynamics tags in the DFT input file to relax all layers of the 2D material and top two layers of the substrate slab.

The workflow has commands for two VASP executables compiled that incorporate vdW-corrections for performing DFT calculations for (1) 2D materials and (2) 3D materials. The first executable is a custom executable to relax 2D materials with a large vacuum and prevent the vacuum from shrinking by not letting the cell length change in the direction of vacuum spacing. The second executable allows the cell volume to change in all directions. Other optional arguments used to initialize the workflow include dipole correction for substrate slabs, tags for database entries, and avenues to modify the INCAR of each firework in the workflow. The parameters *vis* and *vis_ i* where *i*=2d, 3d2d, bulk, trans, and iface are used to override the default *VaspInputSet* with one provided by the user. This can be provided for all fireworks using *vis* or for a specific firework using *vis_ i*. The parameters *uis* and *uis_ i* can be set to change the default settings in the INCAR. The parameter *uis* will set the specified parameters for all INCARs in the workflow, while *uis_ i* will set the INCAR parameters for the corresponding firework.

The first four fireworks in the workflow are optional to reduce computational cost while the final firework generates and performs structural optimization for the 2D-substrate heterostructure configurations. To skip a firework, the custom specification dictionary (*user_ additions*) is used and to determine the ΔE^f_{ads} the system energies and total number of atoms for each skipped firework must be supplied. These energies are stored in the *analysis_ info* dictionary within *user_ additions* to calculate stability related information. The first four structure optimization fireworks may be skipped by setting any of the *is_ i_ optimized* (*i*=bulk, sub, 2d, and 3d2d) flags to $True$ in the *user_ additions* dictionary.

Additionally, the *user_ additions* dictionary can be used to skip computing the ΔE^f_{vac}, ΔE_{b}, or ΔE^f_{ads} energies for the current workflow, specify alternative types of functionals for DFT calculations, and allocate sufficient computing resources to larger slab/heterostructure calculations. The *user_ additions* can add tags to each firework step in the workflow. The *vdw* tag can be used to specify alternative types of functionals. To selectively skip the energetic analysis of the ΔE^f_{vac}, ΔE_{b}, and ΔE^f_{ads} the flags *Formation_ Energy*, *Binding_ Energy*, and *Adsorption_ Energy* can be set to $False$. Additional flags for setting the computational wall-time, cpu-cores, etc. are available in the code documentation with a complete description of their effect on the workflow.

C.4 2D, 3D Form of 2D, and Substrate Slab Surface Information

In generating the 49 workflows, a total of 29 DFT calculations were performed using the FW_1 – FW_3 to optimize the substrate, 2D material, and bulk counterpart

of the 2D material, all of which successfully converged and were stored in a MongoDB database. An additional 24 calculations were performed using FW_4 to generate the surface substrate slab. Of the first four FWs, only one calculation resulted in errors that were not handled by custodian which was FW_4 that generates the Pb (110) substrate slab.

Composition	**Database Ref. ID**	ΔE^f_{vac} (eV/atom)	**Sg.**
MoS_2	c2db id: MoS2-MoS2-NM	0.0705	$P\bar{6}m2$
$1T$-NbO_2	c2db id: NbO2-CdI2-FM	0.3896	$P\bar{3}m1$
$1H$-NbO_2	c2db id: NbO2-MoS2-NM	0.4048	$P\bar{6}m2$
ZnTe	doi: 10.1038/s41699-020-0154-y	0.1763	$P\bar{3}m1$

Table C.1. 2D materials information for all materials studied in this work. The first three 2D materials were obtained from the C2DB (Haastrup et al. 2018) while the last was obtained from S. Torrisi et al. (2020). The database reference ID is the C2DB ID and the DOI number for the structures, ΔE^f_{vac} is the formation energy of the 2D materials and Sg. is the spacegroup of the 2D material.

Composition	**material_id**	$E_{above\ hull}$ (eV/aton)	E_{gap} (eV)	**Sg.**
ZnTe	mp-2176	0.0000	1.078	$F\bar{4}3m$
MoS_2	mp-1018809	0.0010	1.336	$P6_3/mmc$
NbO_2	mp-557057	0.0000	0.300	$I4_1/a$

Table C.2. The materials information for the bulk form of the 2D materials (3D2D) used in this work to calculate the formation energy. The material_id is the Materials Project ID, $E_{above\ hull}$ is the energy above the convex hull of the given 3D2D in eV/atom, is the band gap in eV for the 3D2D, and Sg. is the spacegroup of the 3D2D material.

Composition	**material_id**	$E_{above\ hull}$	**ICSD ID**	**Sg.**	**Surface**
Ag	mp-124	0.0024	604635	$Fm\bar{3}m$	(111)
Al	mp-134	0	53774	$Fm\bar{3}m$	(111)
Au	mp-81	0	611625	$Fm\bar{3}m$	(111)
Cu	mp-30	0	655129	$Fm\bar{3}m$	(111)
Hf	mp-8640	0.0712	41519	$Fm\bar{3}m$	(111)
Ir	mp-101	0	64992	$Fm\bar{3}m$	(111)
Mn	mp-8634	0.0799	41509	$Fm\bar{3}m$	(111)
Nd	mp-159	0.0107	645578	$Fm\bar{3}m$	(111)
Ni	mp-23	0	52265	$Fm\bar{3}m$	(111)
Pb	mp-20483	0	648343	$Fm\bar{3}m$	(110)
Pd	mp-2	0	64918	$Fm\bar{3}m$	(111)
Re	mp-8642	0.0588	41522	$Fm\bar{3}m$	(111)
Rh	mp-74	0	191463	$Fm\bar{3}m$	(110)(111)
Sc	mp-36	0.0464	164103	$Fm\bar{3}m$	(110)
Sr	mp-76	0	76162	$Fm\bar{3}m$	(111)
Ta	mp-50	0	183414	$Im\bar{3}m$	(110)
Te	mp-10654	0.0466	52500	$Pm\bar{3}m$	(111)
Ti	mp-6985	0.0634	168322	$Fm\bar{3}m$	(111)
W	mp-91	0	43421	$Im\bar{3}m$	(110)
V	mp-146	0	171003	$Im\bar{3}m$	(111)
Y	mp-9	0.0275	106221	$Fm\bar{3}m$	(111)
Zr	mp-8635	0.0413	41511	$Fm\bar{3}m$	(111)

Table C.3. The materials information for all substrates simulated this work. The material_id is the Materials Project ID, $E_{above\ hull}$ is the energy above the convex hull of the given substrate in eV/atom, ICSD ID is the inorganic crystal structure database ID, Sg. is the spacegroup of the substrate, and surface is the miller plane of the substrate slab in the 2DMSD.

C.5 Adsorption Energy, Outlier Structures, and Substrate Induced Changes in 2D Films

The adsorption formation energy, shown in Appendix Figure C.2, represent the 2D films that have significant structural distortion or have disintegrated on the

substrate. The state of the adsorbed 2D materials that are classified as significant structural distortions or disintegrated for each 2D material are: 4 2D heterostructure configurations for MoS_2, 12 2D heterostructure configurations for $1T$-NbO_2, 1 2D heterostructure configuration for $1H$-NbO_2, and 5 2D heterostructure configurations for ZnTe.

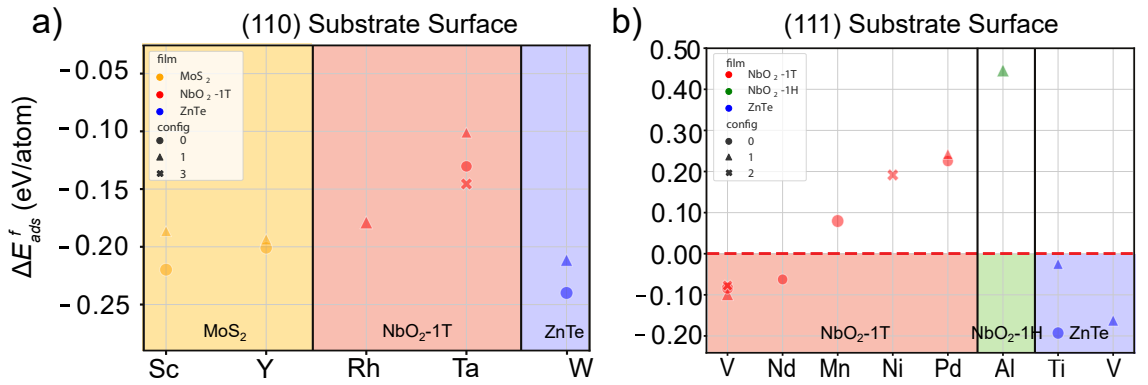

Figure C.2. Adsorption formation energy configurations for the symmetry-matched, low lattice-mismatched (a) (110) and (b) (111) surface substrates where the 2D materials have disintegrated post-adsorption. The (110) surface has more 2D materials that disintegrated than the (111) surface. Each set of symbols (up to 4 points per substrate) represents the unique 2D-substrate configurations.

In Appendix Figure C.2a, the configurations of MoS_2 on Sc (110) surface are buckled with significant disintegration occurring on the substrate surface. The interaction between the (110) Y surface and the S atoms from MoS_2 film interact strongly with some S atoms incorporating into the Y substrate surface resulting in significant distortions of both the 2D material and substrate surface. For $1T$-NbO_2 on the Rh (110) surface, the 2D material disintegrates on the substrate surface. For $1T$-NbO_2 on the Ta (110) surface, the 2D material buckles on the Ta substrate similar to MoS_2 on Ta (110) in Appendix Figure C.4. All configurations for ZnTe on the W (110) surface result in the 2D material disintegrating on the substrate surface.

In Appendix Figure C.2b, all configurations for $1T$-NbO_2 on the V (111) surface have significant disintegration of the 2D material while the substrate surface undergoes some distortions. The circle configuration for the $1T$-NbO_2 on the Nd and Ni (111) surfaces have some degree of structural disorder for the substrate surface and the 2D material has disintegrated. Both configurations for $1T$-NbO_2 on the Pd (111) surface result in disintegration of the 2D film on the substrate surface. For the $1H$-NbO_2 on the Al (111) surface, the 2D material is repelled by the substrate surface resulting in a separation distance of 5.16 Å between the 2D material and substrate surface. For

ZnTe on the Ti and V (111) surface all the 2D materials disintegrate and the Ti (111) surface also has structural disorder.

Appendix Figure C.3 illustrates the change in the thickness of the free-standing 2D material from that of the post-adsorbed 2D material for each 2D-substrate heterostructured material for the structures that were excluded from Chapter 5 Figure 5.6. The thickness of the free-standing/adsorbed 2D material is computed first by finding the average z coordinate of the top and bottom layer of the 2D material given by $\bar{d}_z = \sum_{i=1}^{n} d_{i,z}^{top}/n - \sum_{i=1}^{m} d_{i,z}^{bottom}/m$ where $d_{i,z}$ is the z coordinate of the i^{th} atom summed up to n and m, the total number of atoms in the top and bottom layers, respectively. The thickness, obtained by taking the difference between the average thickness of the adsorbed 2D material from that of the free-standing 2D material, $\delta d = \bar{d}_z^{adsorbed} - \bar{d}_z^{free}$, with positive (negative) values corresponding to an increase (decrease) in the thickness of the adsorbed 2D material.

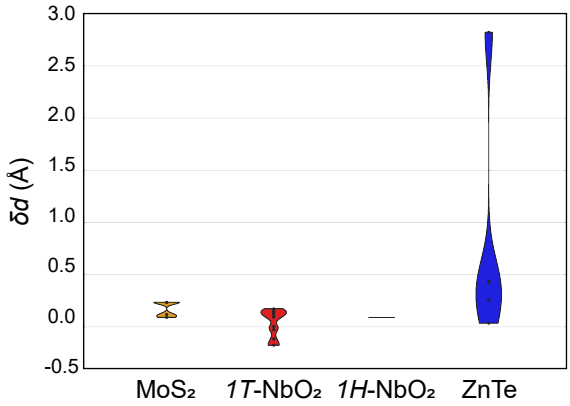

Figure C.3. Each 2D material is separated spatially along the x-axis using a violin plot. The change in the 2D material's thickness, δd, for all substrates is plotted along the y-axis. A positive y-value indicates the 2D material's thickness has increased during adsorption onto the substrate slab. The width of the violin plot is non-quantitative from scaling the density curve by the number of counts per violin, however, within one violin plot, the relative x-width does represent the frequency that a 2D material's thickness changes by y amount relative to the total number of data points in the plot. These structures are consistent with Appendix Figure C.2.

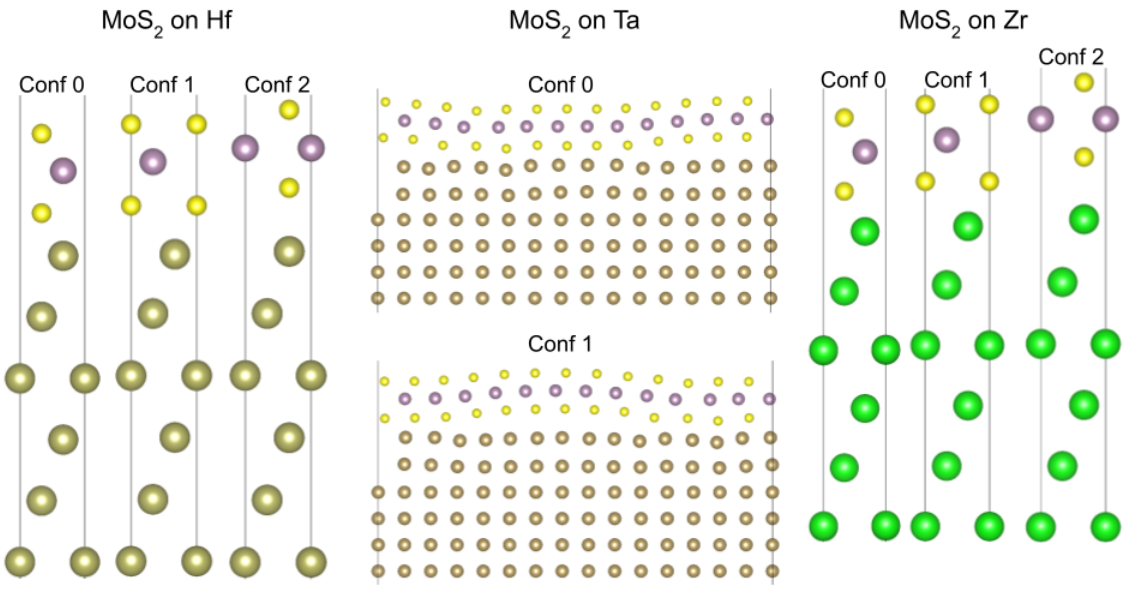

Figure C.4. Structure models for 2D-substrate heterostructures in Chapter 5 that are discussed as the adsorption formation energy outliers. The 2D-substrate heterostructure configurations, from left to right, are MoS_2 on Hf (111), Ta (110), and (111) Zr.

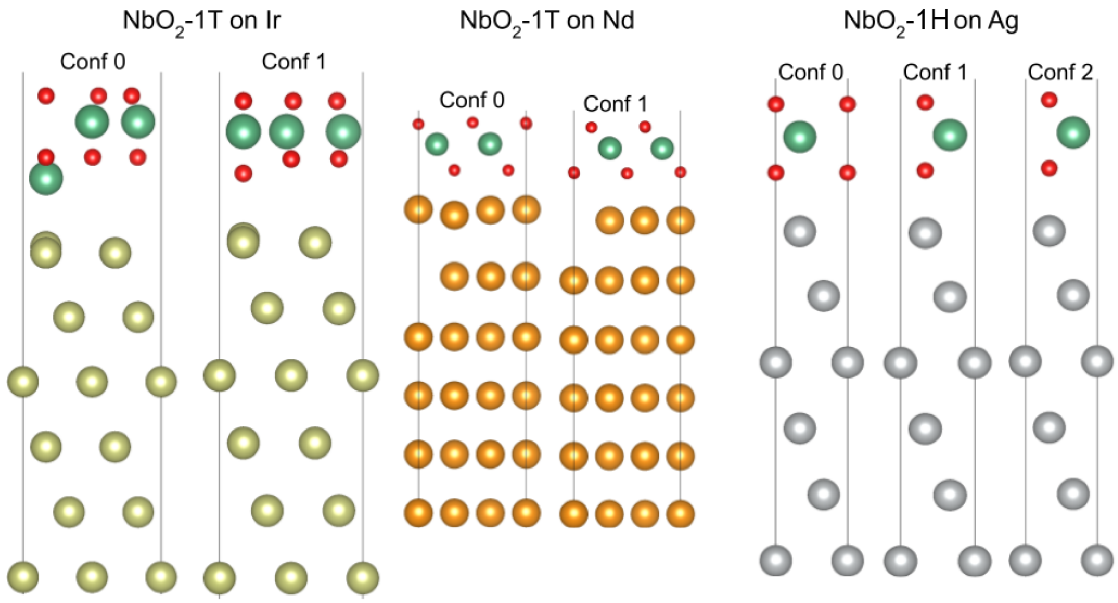

Figure C.5. Structure models for 2D-substrate heterostructures in Chapter 5 that are discussed as the adsorption formation energy outliers. The 2D-substrate heterostructure configurations, from left to right, are $1T$-NbO_2 on Ir (111), $1T$-NbO_2 on Nd (111) surface, and $1H$-NbO_2 on Ag (111) surface.

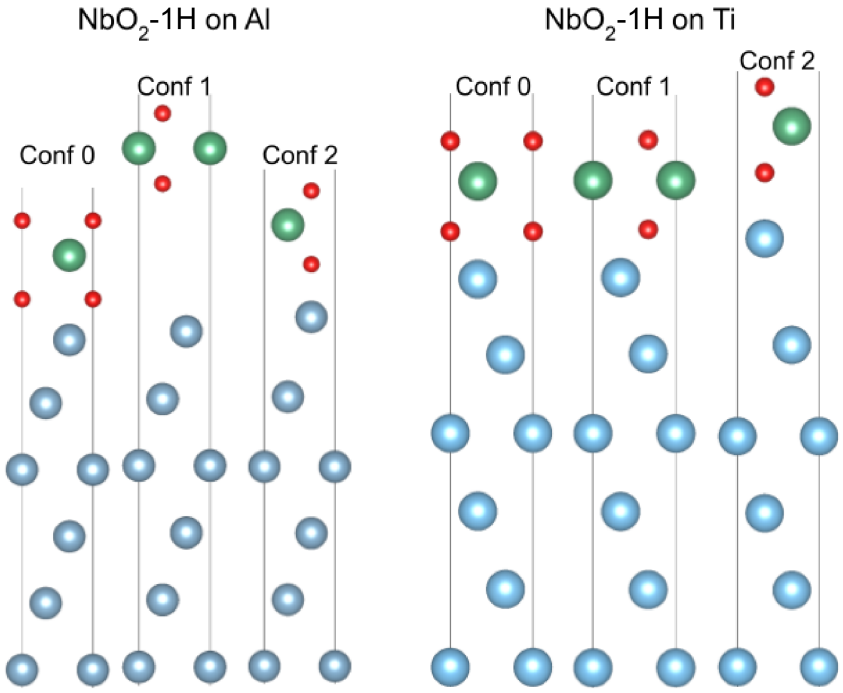

Figure C.6. Structure models for 2D-substrate heterostructures in Chapter 5 that are discussed as the adsorption formation energy outliers. The 2D-substrate heterostructure configurations, from left to right, are $1H$-NbO_2 on Al and Ti (111) surface.

APPENDIX D

SUPPLEMENTARY MATERIAL FOR CHAPTER 6

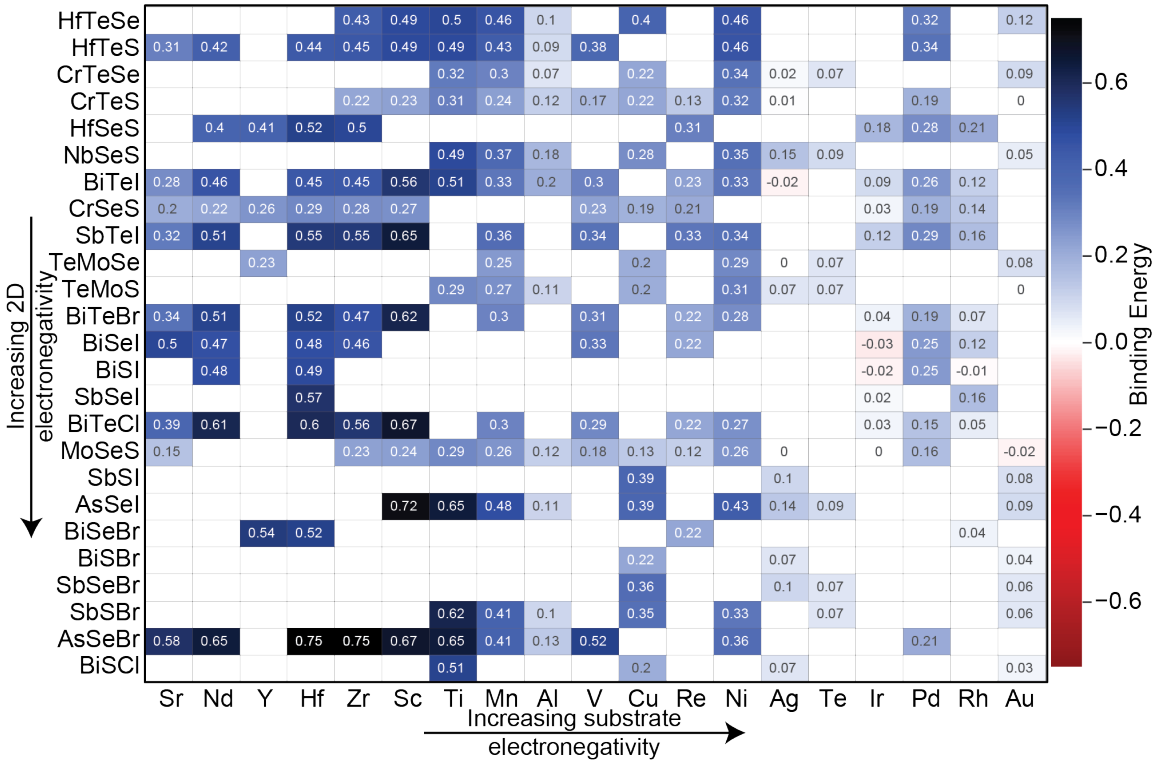

Figure D.1. Heatmap for the 1H phase of the Janus 2D materials sorted from least (top/left) to greatest (bottom/right) electronegativity of the 2D materials and substrate. The colormap represents the (ΔE_{b}) binding energy.

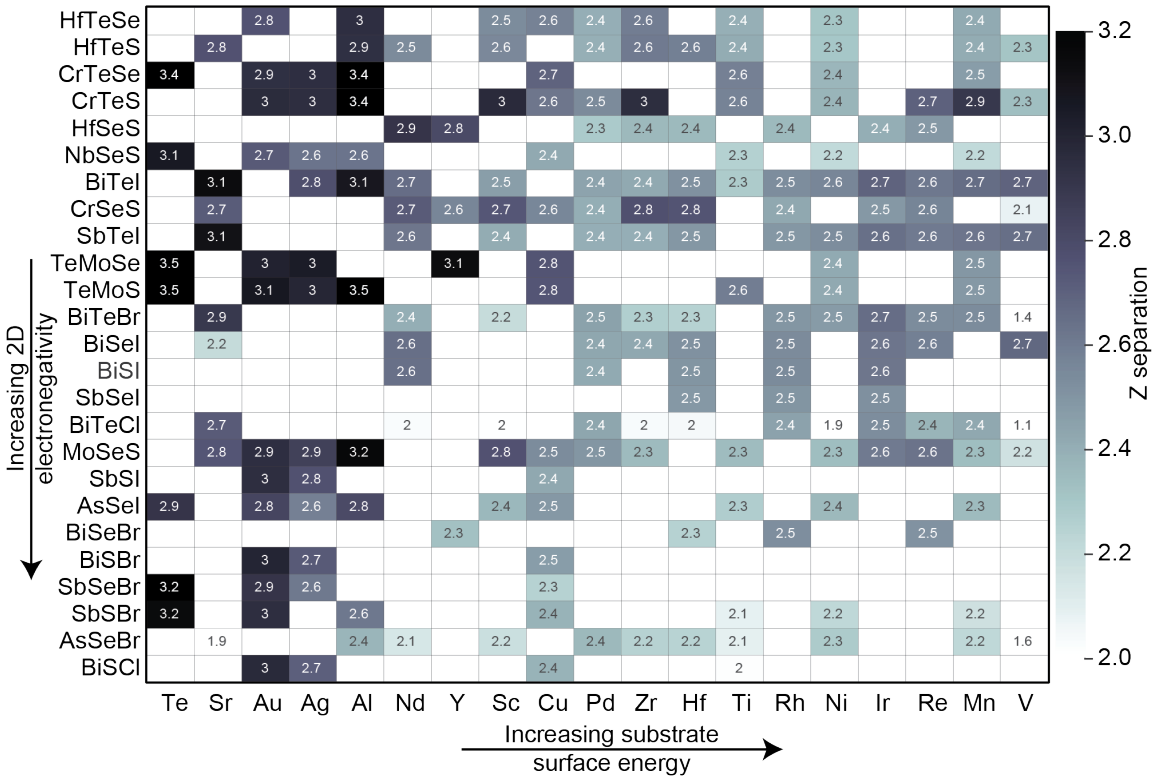

Figure D.2. Heatmap for the $1H$ phase of the Janus 2D materials sorted from least (top/left) to greatest (bottom/right) electronegativity of the 2D materials and substrate's surface energy. The colormap represents the z-separation.

D.1 Constituent 2D-substrate Heterostructure Information

The unique elemental composition of Janus 2D materials results in few 3D phases that possess all constituent elements and very few 3D phases which have a direct composition match to the Janus 2D material. However, a composition matched bulk phase for some of the Janus 2D materials can be obtained through a mixture of two layered bulk counterparts. In fact, current synthesis methods for Janus 2D phases use the parent classical layers (i.e. MoS_2 and $MoSe_2$) to synthesize the Janus 2D material (MoSeS) Qin et al. 2022. To access more Janus 2D phases, the formation energy (ΔE^f_{vac}) for 2D materials without a corresponding composition matched 3D phase is computed using a mixture of the layered bulk counterparts. The information for each bulk phase used to compute the formation energy is listed in Table D.2 and the corresponding Janus 2D phases are listed in Table D.1. Table D.3 provides additional details regarding the substrates used in this study while Table D.4 lists all lattice-matched substrates for each 2D material.

Composition	**Polytype**	ΔE^f_{vac} (eV/atom)	**c2db_id**	**Sg.**
AsSeBr	T	-0.699	AsBrSe-989f469f06bd	156
AsSeBr	H	-0.552	AsBrSe-206b9dcf2af6	156
AsSeI	T	0.109	AsISe-5d829e480507	156
AsSeI	H	0.265	AsISe-ca926a42865b	156
BiSBr	T	0.117	BiBrS-49b7be14f786	156
BiSBr	H	0.226	BiBrS-3b305c3e2c18	156
BiSCl	H	0.23	BiClS-99fd027b1d0b	156
BiSCl	T	0.116	BiClS-c96ef4fc869c	156
BiSI	T	0.121	BiIS-acdcd16c0d76	156
BiSI	H	0.239	BiIS-40034665f9f1	156
BiSeBr	T	0.084	BiBrSe-de5756e4fbfa	156
BiSeBr	H	0.189	BiBrSe-11db0908d9ef	156
BiSeCl	T	0.094	BiClSe-a80866a2c6b4	156
BiSeI	H	0.199	BiISe-433f707c632c	156
BiSeI	T	0.092	BiISe-70cbc0e44d36	156
BiTeBr	H	0.218	BiBrTe-f4f45fcade85	156
BiTeBr	T	0.108	BiBrTe-304bc6a92d82	156
BiTeCl	H	0.229	BiClTe-badda86cab42	156
BiTeCl	T	0.104	BiClTe-968a6902b7f5	156
BiTeI	H	0.21	BiITe-a84d988e38ac	156
BiTeI	T	0.107	BiITe-2d41b3dd1772	156
CrSeS	H	0.122	CrSSe-09e1e5ef94cb	156
CrSeS	T	0.262	CrSSe-370baf5c3264	156
CrTeS	T	0.211	CrSTe-7aa3d966f272	156
CrTeS	H	0.149	CrSTe-8a0864d30ce1	156
CrTeSe	T	0.449	CrSeTe-076280734e10	156
CrTeSe	H	0.359	CrSeTe-5d9d3ded04de	156
HfSeS	H	0.268	HfSSe-9afb20358166	156
HfSeS	T	0.088	HfSSe-63618e5bf062	156
HfTeS	T	0.093	HfSTe-6da5c5b7dd23	156
HfTeS	H	0.234	HfSTe-2602a918955b	156
HfTeSe	H	0.226	HfSeTe-305c779b8752	156
HfTeSe	T	0.099	HfSeTe-a514ac4ca101	156
MoSeS	T	0.506	MoSSe-6a7e9eb85b12	156
MoSeS	H	0.271	MoSSe-de7ac5fc6945	156
NbSeS	H	0.901	NbSSe-596b8fdeb3d4	156
NbSeS	T	0.943	NbSSe-91a283a2d283	156
SbSBr	H	0.272	BrSSb-4ae37f15e1fe	156

SbSBr	T	0.152	BrSSb-4da5c6be60db	156
SbSI	T	0.16	ISSb-5b94060698bc	156
SbSI	H	0.292	ISSb-4c49d27e66e5	156
SbSeBr	T	0.113	BrSbSe-89b15ddef41d	156
SbSeBr	H	0.228	BrSbSe-c2a344b393f0	156
SbSeI	T	0.122	ISbSe-df0019ec24b5	156
SbSeI	H	0.241	ISbSe-343d2125478e	156
SbTeI	H	0.207	ISbTe-052a3116531d	156
SbTeI	T	0.095	ISbTe-0f02957b17cf	156
TeMoS	T	0.187	MoSTe-2ea941c8bc3c	156
TeMoS	H	0.032	MoSTe-e4bb8738150a	156
TeMoSe	T	0.07	MoSeTe-f3f66ff2fdc2	156
TeMoSe	H	-0.101	MoSeTe-42eb12e7b656	156

Table D.1. Two-dimensional (2D) materials information for all materials studied in this work. All Janus 2D materials are obtained from the C2DB Haastrup et al. 2018. The composition, Materials Project ID, structural polytype, the computed formation energy in eV/atom, the C2DB ID, and the spacegroup number of each Janus material.

Composition	**material_id**	**Sg.**
AsSeBr	mp-631257	216
AsSeI	mp-505373	14
BiSBr	mp-23324	62
BiSCl	mp-23318	62
BiSI	mp-23514	62
BiSeBr	mp-569707	62
BiSeCl	mp-610491	62
BiSeI	mp-23020	62
BiTeBr	mp-33723	156
BiTeCl	mp-28944	186
BiTeI	mp-22965	156
Cr3(TeSe)2	mp-1226346	12
Cr4(SeS)3	mp-1226292	146
Cr4Te3S	mp-1226328	44
HfSeS	mp-1224282	156
HfTeS	mp-1224272	156
HfTeSe	mp-1224271	156
Mo2SeS3	mp-1221485	156
Nb5(SeS)2	mp-1220382	12
SbSBr	mp-22971	62
SbSI	mp-23041	62
SbSeBr	mp-1209072	62
SbSeI	mp-22996	62
SbTeI	mp-28051	12
Te3Mo2Se	mp-1217371	156
TeMo2S3	mp-1217334	156

Table D.2. The corresponding bulk (3D) phases used to determine the formation energy (ΔE^f_{vac}) of the Janus 2D materials. The composition, Materials Project ID, and the spacegroup number for each 3D phase.

Composition	**material_id**	**Surface Energy** (eV/Å^2)	**Sg.**
Ag	mp-124	0.0472	225
Al	mp-134	0.0474	225
Au	mp-81	0.0445	225
Cu	mp-30	0.0966	225
Hf	mp-8640	0.1234	225
Ir	mp-101	0.151	225
Mn	mp-8634	0.1827	225
Nd	mp-159	0.0547	225
Ni	mp-23	0.1402	225
Pd	mp-2	0.097	225
Re	mp-8642	0.1653	225
Rh	mp-74	0.1303	225
Sc	mp-36	0.0838	225
Sr	mp-76	0.0288	225
Te	mp-10654	0.0009	221
Ti	mp-6985	0.128	225
V	mp-146	0.1913	229
Y	mp-9	0.0683	225
Zr	mp-8635	0.113	225

Table D.3. The materials information for the bulk substrate slabs. The composition, Materials Project ID, the computed surface energy for the (111) in eV/Å^2, and the spacegroup number of each substrate material used in this study.

2D Composition	Substrates
AsBrSe-206b9dcf2af6	Al Hf Mn Nd Ni Pd Sc Sr Ti V Zr
AsBrSe-989f469f06bd	Al Au Cu Mn Ni Pd Sc Te Ti
AsISe-5d829e480507	Au Cu Te Ti
AsISe-ca926a42865b	Ag Al Au Cu Mn Ni Sc Te Ti
BiBrS-3b305c3e2c18	Ag Au Cu
BiBrS-49b7be14f786	Hf Ir Nd Pd Rh
BiBrSe-11db0908d9ef	Hf Re Rh Y
BiBrSe-de5756e4fbfa	Hf Ir Nd Pd Re Rh Sc Sr V Zr
BiBrTe-304bc6a92d82	Ag Al Cu Ir Mn Ni Pd Re Sc Sr Ti V Zr
BiBrTe-f4f45fcade85	Hf Ir Mn Nd Ni Pd Re Rh Sc Sr V Zr
BiClS-99fd027b1d0b	Ag Au Cu Ti
BiClS-c96ef4fc869c	Hf Rh
BiClSe-a80866a2c6b4	Hf Ir Nd Pd Re Rh Sr V Y Zr
BiClTe-968a6902b7f5	Ag Al Hf Ir Mn Nd Ni Pd Re Rh Sc Sr Ti V Zr
BiClTe-badda86cab42	Hf Ir Mn Nd Ni Pd Re Rh Sc Sr V Zr
BiIS-40034665f9f1	Hf Ir Nd Pd Rh
BiIS-acdcd16c0d76	Hf Ir Mn Nd Ni Pd Re Rh Sc Sr V Zr
BiISe-433f707c632c	Hf Ir Nd Pd Re Rh Sr V Zr
BiISe-70cbc0e44d36	Al Hf Ir Mn Nd Ni Pd Re Rh Sc Sr Ti V Zr
BiITe-2d41b3dd1772	Ag Al Au Cu Mn Ni Sc Te Ti
BiITe-a84d988e38ac	Ag Al Hf Ir Mn Nd Ni Pd Re Rh Sc Sr Ti V Zr
BrSSb-4ae37f15e1fe	Al Au Cu Mn Ni Te Ti
BrSSb-4da5c6be60db	Au Cu Rh Te
BrSbSe-89b15ddef41d	Hf Re Rh
BrSbSe-c2a344b393f0	Ag Au Cu Te
CrSSe-09e1e5ef94cb	Cu Hf Ir Nd Pd Re Rh Sc Sr V Y Zr
CrSSe-370baf5c3264	Cu Hf Ir Mn Nd Pd Re Rh Sc Sr V Y Zr
CrSTe-7aa3d966f272	Ag Al Au Cu Mn Ni Te Ti
CrSTe-8a0864d30ce1	Ag Al Au Cu Mn Ni Pd Re Sc Ti V Zr
CrSeTe-076280734e10	Ag Al Au Cu Mn Ni Te Ti
CrSeTe-5d9d3ded04de	Ag Al Au Cu Mn Ni Te Ti
HfSSe-63618e5bf062	Al Hf Mn Nd Ni Pd Sc Sr Ti V Zr
HfSSe-9afb20358166	Hf Ir Nd Pd Re Rh Y Zr
HfSTe-2602a918955b	Al Hf Mn Nd Ni Pd Sc Sr Ti V Zr
HfSTe-6da5c5b7dd23	Al Cu Mn Ni Pd Sc Ti V Zr
HfSeTe-305c779b8752	Al Au Cu Mn Ni Pd Sc Ti Zr
HfSeTe-a514ac4ca101	Ag Al Au Cu Mn Ni Ti
ISSb-4c49d27e66e5	Ag Au Cu

ISSb-5b94060698bc	Hf Ir Nd Pd Rh Y
ISbSe-343d2125478e	Hf Ir Rh
ISbSe-df0019ec24b5	Hf Ir Nd Pd Re Rh Sc Sr V Y Zr
ISbTe-052a3116531d	Hf Ir Mn Nd Ni Pd Re Rh Sc Sr V Zr
ISbTe-0f02957b17cf	Ag Al Mn Ni Pd Re Sc Sr Ti V Zr
MoSSe-6a7e9eb85b12	Ag Al Hf Ir Mn Ni Pd Re Sc Sr Ti V Zr
MoSSe-de7ac5fc6945	Ag Al Au Cu Ir Mn Ni Pd Re Sc Sr Ti V Zr
MoSTe-2ea941c8bc3c	Ag Al Au Cu Mn Ni Te Ti
MoSTe-e4bb8738150a	Ag Al Au Cu Mn Ni Te Ti
MoSeTe-42eb12e7b656	Ag Au Cu Mn Ni Te Y
MoSeTe-f3f66ff2fdc2	Ag Al Au Cu Mn Ni Te Ti
NbSSe-596b8fdeb3d4	Ag Al Au Cu Mn Ni Te Ti
NbSSe-91a283a2d283	Ag Au Cu Te Y

Table D.4. All Janus 2D materials are listed under composition. The numbers listed after the '-' are the associated C2DB ID for the 2D material. The substrates column lists the corresponding lattice-matched substrates identified given the heterostructure matching criteria chosen in this work for the given Janus 2D material.

D.2 Electronic Structure and Charge Density Benchmark Calculations

In order to gauge the accuracy (significant figures) to which a particular computed value can be determined given the default settings in the *CMDLElectronic* set convergence, tests are performed using the various grid densities for the charge density and density of states. The relevant parameters for the density of state are typically the accuracy of the (a) valence band maximum (VBM), (b) conduction band minimum (CBM), (c) Fermi level, and (d) band gap. For Bader analysis, the total charge or number of electrons associated with each atom is of importance. Various partition methods exist in literature (Choudhuri and Truhlar 2020) to compute the total charge for atoms. Bader analysis uses zero flux surfaces or local minimums in the charge density surface to partition atoms into volumes and approximate the charge on an atom by summing the number of electrons enclosed within that region (Henkelman, Arnaldsson, and Jønsson 2006). It is important to note that different methods for partitioning the charge density and atomic volume has a significant impact on the total charge assigned to each atom. As described in the work by Choudhuri and Truhlar (2020) the variations could be greater than 1 atomic charge unit and depend on method as well as the system of interest along with, to a lesser extent, the choice of pseudopotential.

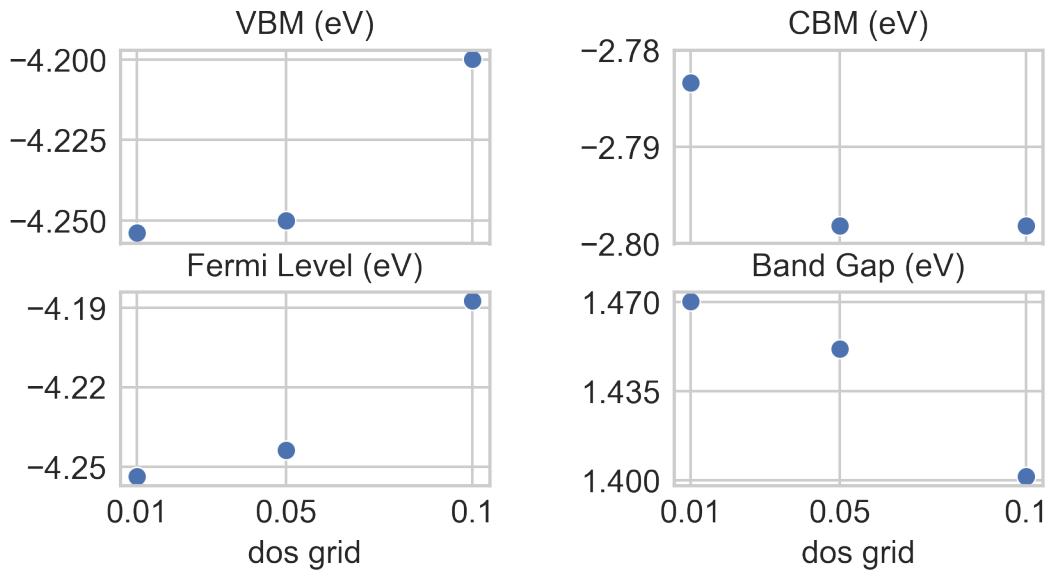

Figure D.3. The (a) valence band maximum (VBM), (b) conduction band minimum, (c) Fermi level, and (d) band gap for density of states spacings of 0.1, 0.05, and 0.01 eV. The difference between the 0.01 and 0.05 eV grid spacings are very small for each quantity.

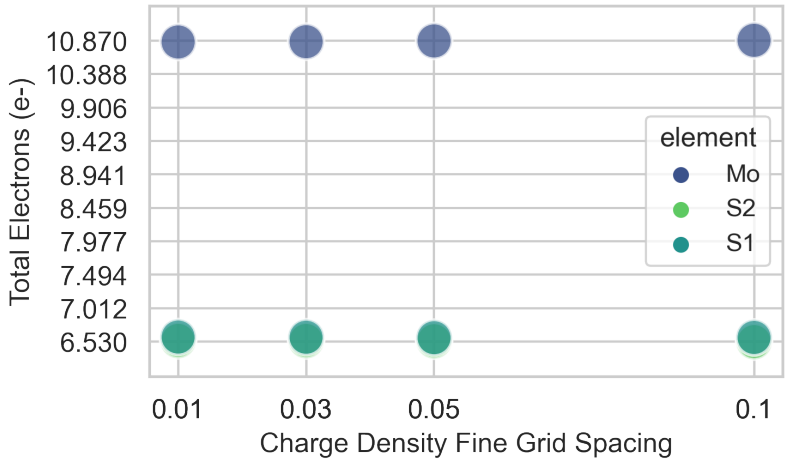

Figure D.4. The Bader charge on each atom given four fine grid spacings (NG{X,Y,Z}G): 0.1, 0.05, 0.03, and 0.01. The total charges on each atom have near zero variation (0.004 electrons per atom). The charges on the respective atoms are sufficiently converged given the choice of grid spacing.

The values described above are highly dependant on the grid spacings; as such it

is essential to determine how variations in the grid densities impact these quantities. Appendix Figure D.3 provides benchmark calculations performed by varying the number of points used to calculate the density of states, while Appendix Figure D.4 illustrates variations in the Bader charges with variations in the charge density grid. The density of states are well converged given the default grid spacing of 0.05 eV resulting in minimal changes in the VBM, CBM, Fermi level, and band gap. The band gap has the largest variation of 0.019 eV between the finer grid spacing of 0.01 to 0.05. The Bader charges show convergence to 0.004 electrons given the chosen grid choice of 0.03 compared to the 0.01 grid spacing. To further confirm the charge density does not result in significant changes with increasing grid density, Appendix Figure D.5 plots the z-projected charge density at each grid spacing. The charge density does not display any discernible differences in this plot.

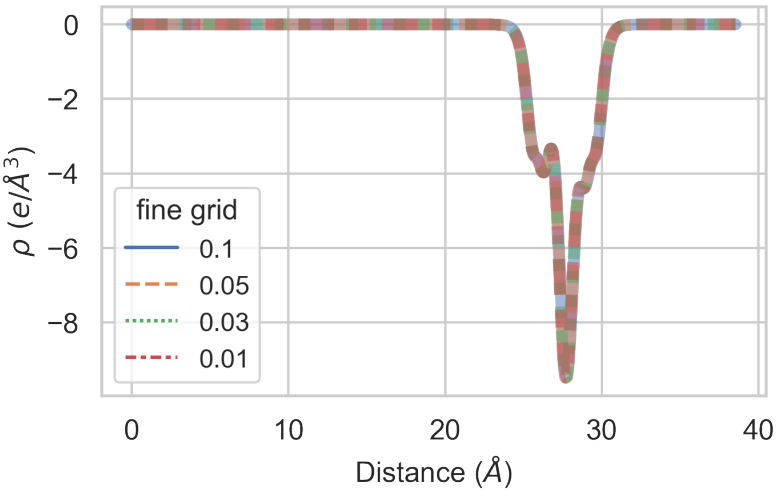

Figure D.5. The z-projected charge density at four different fine grid spacings (NG{X,Y,Z}G): 0.1, 0.05, 0.03, and 0.01. There is no discernible difference between the charge density line profiles for any of the grid spacings.

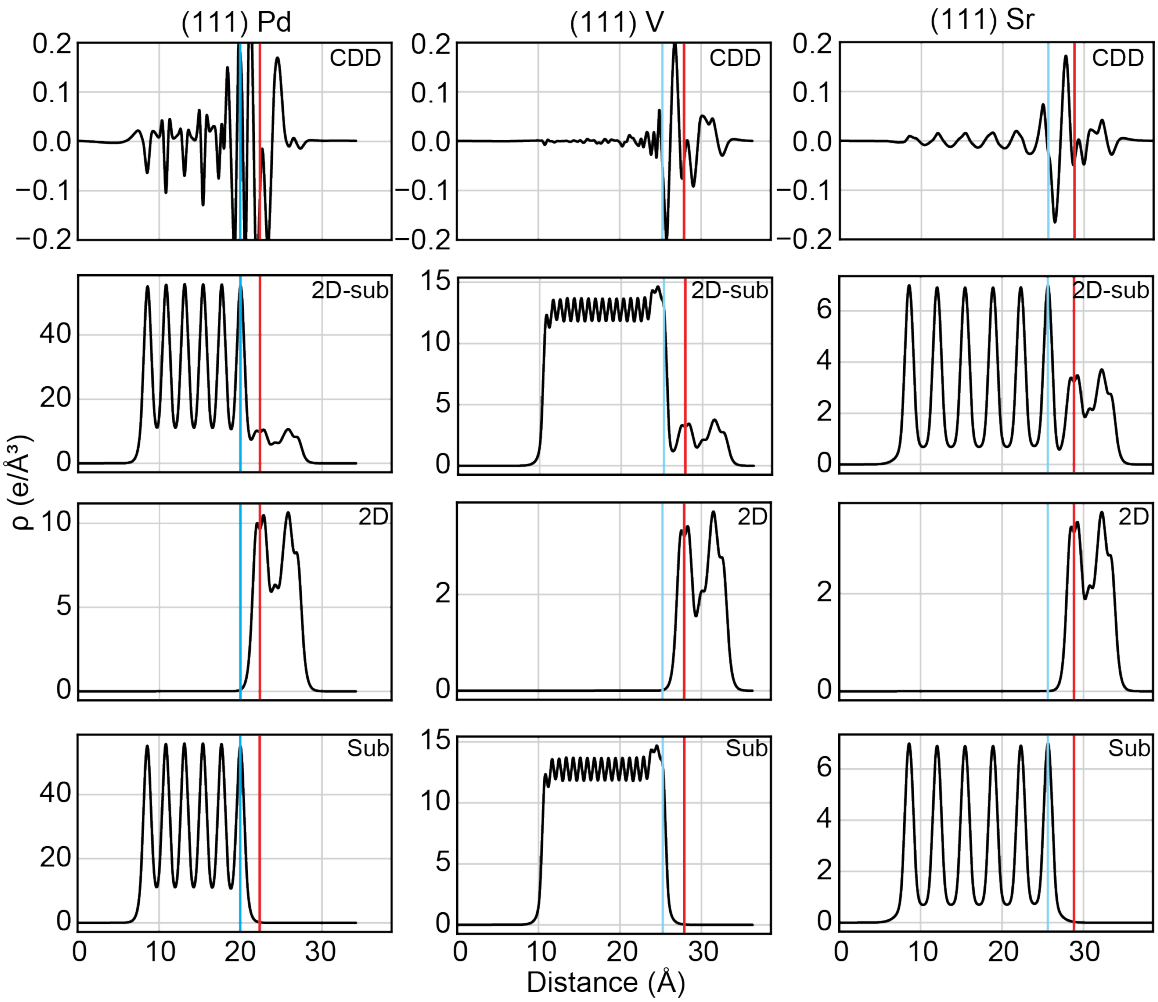

Figure D.6. From top to bottom the z projected charge density difference, the charge density of the heterostructure, isolated 2D material, isolated substrate charge density. The blue line represents the average z position of the substrate and the red line represents the average z position of the bottom layer of the 2D material.

D.3 Machine Learning Convergence Testing

To increase the speed and accuracy of the random forest models, feature selection and convergence testing is performed to evaluate the models robustness. The first step in the process is removing features in the data set with values that occur frequently, thus producing a low-variance feature. These features are unlikely to be useful in determining the target property and can significantly decrease the speed of training the model, as well as increase the computational time spent finding optimal hyper-parameters.

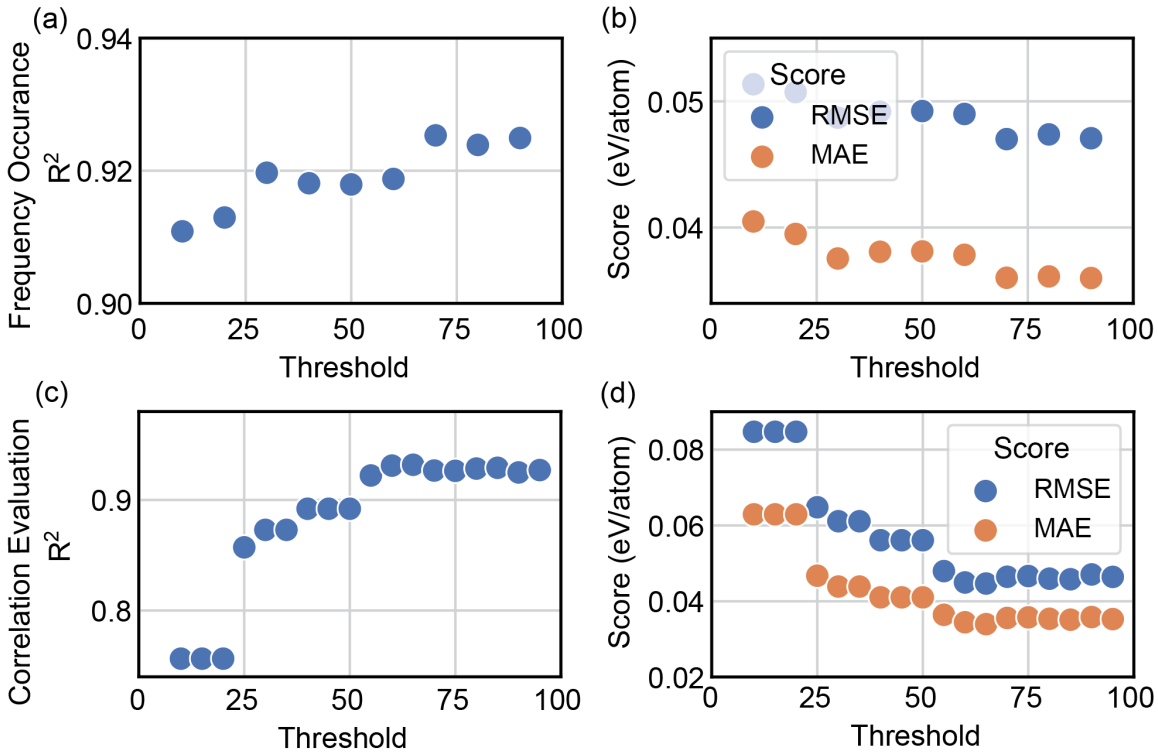

Figure D.7. The binding energy random forest regression model's performance metrics measured using the (a, c) R^2 value and (b, d) the model accuracy evaluated with the RMSE and MAE. For each threshold value in (a, b), features which have a maximum frequency occurrence of any value over the threshold are removed from the training set. For each threshold value in (c, d), features which have a correlation percentage of any value over the threshold are removed from the training set.

The starting number of features in the data set is 155 features. For each feature, the maximum occurrence (frequency) for each unique numeric value is computed. For example, if a feature contains 4 values of X and 1 value of Y, the feature would have a maximum frequency of 4/5 or 80%. The new data set is created by removing all features above a certain threshold value. To find the optimal threshold value, features are removed over a range of values from 10-90 in increments of 10. A test model is then used to benchmark the model's performance for each threshold value and the R^2, RMSE, and MAE values are plotted. The test model was set to the default parameters in *scikit-learn* and K=5 for cross-validation. Appendix Figure D.7a-Figure D.7b show the test model's performance in predicting the binding energy for each threshold value. The best results are obtained when dropping features with a percent occurrence greater than 80%. This results in 31 features removed for the binding energy machine learning model.

The next feature selection method implemented was to remove highly correlated

features. The starting number of features was 124 features. The correlation between features was determined using *pandas* correlation matrix function *corr*. Threshold values from 10%-95% were searched over successively to find the optimal value which improves or maintains the test model's performance. A threshold value of 65% was found to improve the model performance which was measured using the R^2, RMSE, and MAE for the range of threshold values shown in Appendix Figure D.7c-Figure D.7d. The number of highly correlated features removed was 102, leaving 22 features remaining.

The final method to remove unnecessary features was the recursive feature elimination with cross-validation (RFECV) method implemented in *scikit-learn*. This feature selection method recursively removes one feature at a time from the feature space, trains and tests the model's performance returning the optimal features that maximizes the model's performance. A total of 6 features were removed using this method leaving 16 features for training and testing the model. The 16 features are: substrate's surface energy (eV/Å^2), ΔE^f_{vac} (eV/atom), C2DB heat of formation (eV/atom), C2DB (eV/atom), atomic number of the 2D-M, -X, and substrate, average ionic radii of the 2D-M, average electronegativity of the 2D and substrate, z-separation distance, maximum packing efficiency of the 2D and substrate, range of the constituent element's molar volume, number of symmetry operation of the heterostructure, and average number of s valence electrons in the heterostructure. If the element does not have an average ionic radius implemented in *pymatgen*, the atomic radii is used instead.

The same procedure discussed above was implemented for the z-separation distance machine learning model. Appendix Figure D.8 shows the performance of the model for the removal of low-variance and strongly correlated features. The threshold values obtained for the binding energy model were also sufficient for the z-separation model. Thus, the same features were removed in the first 2 features selection steps for both models. For clarity, only the RMSE values are shown in Appendix Figure D.8b and Appendix Figure D.8d as the absolute values for the RMSE and MAE are an order of magnitude smaller. However, the trend is identical for both metrics and provides no additional information about the models performance that cannot be obtained from the RMSE plot. Starting with 22 features, the RFECV method removed a total of 11 features leaving 11 features for training and testing the model. The 11 features are: substrate's surface energy (eV/Å^2), ΔE^f_{vac} (eV/atom), C2DB heat of formation (eV/atom), atomic number of the 2D-X and substrate, average electronegativity of the substrate, change in thickness of the adsorbed 2D material, maximum packing efficiency of the 2D and substrate, range of the constituent element's molar volume, and the average number of s valence electrons in the heterostructure.

The hyper-parameter search for the final random forest regression model was performed using *scikit-learn*'s *GridSearchCV* method. The hyper-parameters searched

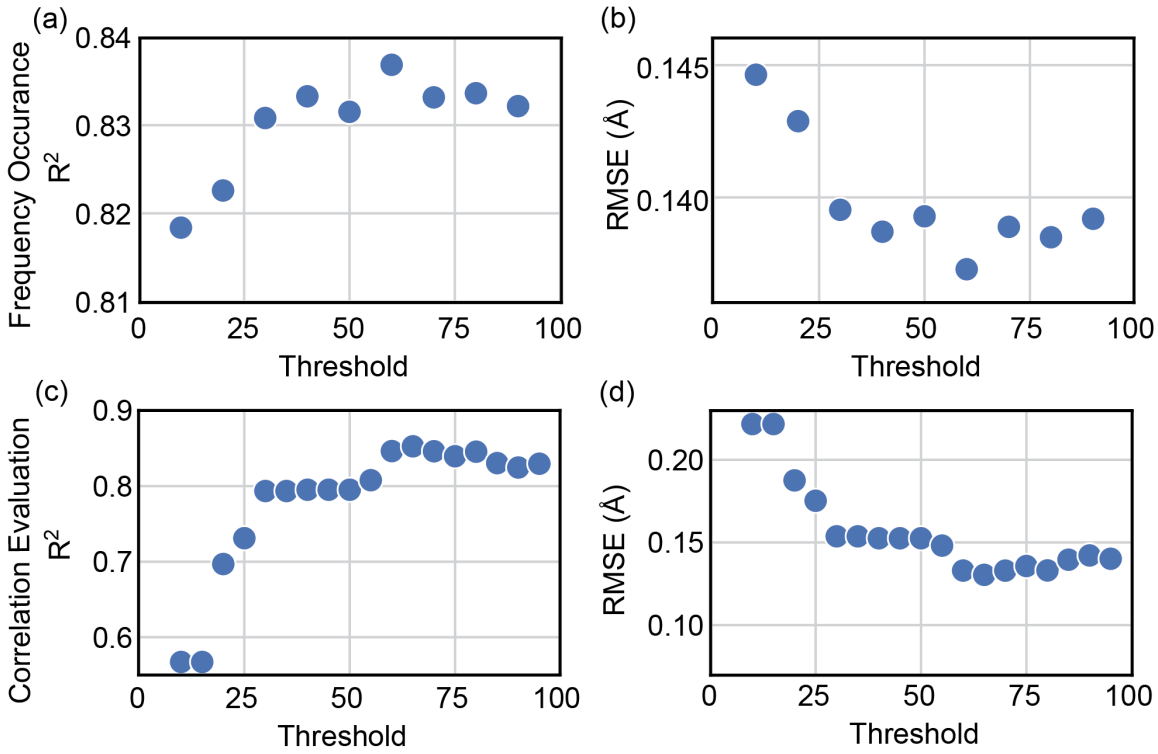

Figure D.8. The z-separation random forest regression model's performance metrics. The (a, c) R^2 value and (b, d) the model accuracy evaluated with RMSE. For each threshold value in (a, b), features which have a maximum frequency occurrence over the threshold are removed from the training set. For each threshold value in (c, d), features which have a correlation percentage of any value over the threshold are removed from the training set.

over for the binding energy model were: number of trees (n_estimators)=[50, 100, 150, 200, 250, 300, 350, 400, 450, 500], the maximum number of features to consider when looking for the best split (max_features)=[auto, sqrt, log2], the maximum depth of the tree (max_depth)=[10, 20, 30, 40, 50, 60, 70, 80, 90, 100, 110, None], the minimum number of samples required to split an internal node (min_samples_split)=[2, 3, 4, 5], the minimum number of samples required to be at a leaf node (min_samples_leaf)=[1, 2, 4], and the method of selecting samples for training each tree (bootstrap)=[True, False]. The optimal hyper-parameters were: n_estimators=150, max_features=sqrt, max_depth=50, min_samples_split=2, min_samples_leaf=1, and bootstrap=False.

The hyper-parameters searched over for the z-separation distance model were: number of trees (n_estimators)=[50, 100, 150, 200, 250, 300, 350, 400, 450, 500], the maximum number of features to consider when looking for the best split (max_features)=[auto, sqrt, log2], the maximum depth of the tree (max_depth)=[50, 75, 100, 125, 150, 175, 200, None], the minimum number of samples required to

split an internal node (min_samples_split)=[2, 3, 4, 5], the minimum number of samples required to be at a leaf node (min_samples_leaf)=[1, 2, 4], and the method of selecting samples for training each tree (bootstrap)=[True, False]. The optimal hyper-parameters were: n_estimators=500, max_features=sqrt, max_depth=50, min_samples_split=2, min_samples_leaf=1, and bootstrap=False.

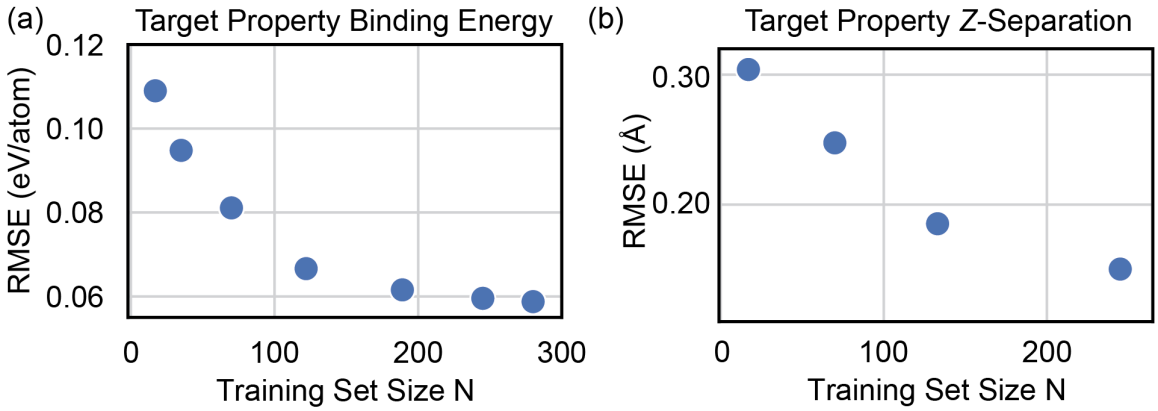

Figure D.9. The RMSE score verses the training set size N for predicting the (a) binding energy and (b) z-separation distance. The accuracy for predicting the binding energy converges around a sample set size of 200 while the training sample size for the z-separation distance converges closer to 300 samples in the training set.