

Towards Efficient Computational Predictions of Battery Cathodes

by

Gregory Houchins

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

at

Carnegie Mellon University
Department of Physics
Pittsburgh, Pennsylvania

Advised by Professor Venkat Viwanathan

August 23, 2020

Abstract

A series of computational tools were employed with the primary goal of understanding layered transition metal oxide materials as lithium ion battery cathodes. As nearly all of the data within this thesis was generated from density functional theory (DFT), we begin with an analysis of the uncertainty of DFT with respect to the choice of exchange correlation functional through the development of a prediction confidence metric and the propagation of error through the Debye-Grüneisen model for lattice vibrations. The prediction confidence metric is applied to the study of transition metal ordering in layered Ni-Mn-Co (NMC) oxide cathodes and enables us to rationalize the disagreement with experimentally seen phases. For the purpose of accelerating the computational predictions of these materials, we then train a neural network potential for the prediction of energy and forces using atom centered symmetry functions as the featurization. The success of this highly accurate machine learning potential is seen through its ability to recreate the thermodynamic properties with an added error that is below the error of the underlying DFT itself. We then predict the open circuit voltage for a series of NMC compositions as well as the lattice dynamics during cycling that have been linked to degradation of the cathode. We then quickly explore a promising machine learning algorithm that is beyond the fingerprint based methods conventionally used. Finally, we dive deeper into the mechanism of another avenue of degradation in the release of highly reactive singlet oxygen seen in NMC, as well as Li-air and Na-air batteries. We provide a unified picture for the mechanism, effect of electrolyte properties, and onset potential for this singlet oxygen release.

Acknowledgments

I would like to first thank my advisor, Venkat Viswanathan who has both allowed and encouraged the large amount of intellectual freedom during my graduate work. It has been exciting and an honor to be a part of the tremendous growth within his group, both in size and in research strengths.

I would also like to thank the many colleagues and collaborators for aiding in the work within this thesis. I would like to thank Holden Parks for his assistance in running and explaining the phonon calculations with the machine learning potentials, Pinwen Guan for the collaboration on the Debye model, Erik Thiede for his significant contribution to in the covariant convolutional neural network work, and Vikram Pande for his contribution to the work on singlet oxygen.

I also would like to thank some past and current members of the group that I did not have the honor of collaborating with but with whom I have shared both friendship and many insightful scientific discussions, including Adarsh Dave, Abhishek Khetan, and Victor Venturi.

And finally, I would like to give a huge thanks to Maya who not only has given me invaluable love and support, but helped to improve the visuals within this thesis.

Thanks you.

Contents

1	Introduction	1
1.1	Overview of Thesis and Contributions	2
1.1.1	Uncertainty Quantification	2
1.1.2	Voltage Prediction	3
1.1.3	Battery Degradation	3
1.2	Organization of Thesis	3
2	Methods	5
2.1	Density Functional Theory	5
2.1.1	Kohn-Sham Equations	6
2.1.2	Exchange Correlation Functionals	9
2.2	Machine Learning Potentials	10
2.2.1	Atom Centered Symmetry Functions	10
2.3	Debye-Grüneisen Model	14
2.3.1	Original Debye Model	14
2.3.2	Anharmonic Corrections	17
3	Uncertainty Quantification in Density Functional Theory	20
3.1	Bayesian Error Estimation Functional	21
3.2	Prediction Confidence	25
3.3	Uncertainty Quantification in Thermodynamic Predictions	32
3.3.1	Application of the Debye-Grüneisen Model	33
3.4	Summary	37
4	Transition Metal Cathodes	38
4.1	Cluster Expansion	44
4.1.1	Training Set	47
4.1.2	Model Selection	47
4.1.3	Prediction of Transition Metal Ordering	49
4.1.4	Prediction of Average Voltage	56
4.2	Atom Centered Symmetry Functions	58
4.2.1	Hyperparameter Selection	59
4.2.2	Thermodynamic Properties	62

4.2.3	Open Circuit Voltage	66
4.2.4	Structural Changes During Cycling	70
4.3	Summary	75
5	Advanced Machine Learning Potentials	77
5.1	Review of Current Methods	77
5.2	Covariant Convolutional Neural Networks	79
6	Singlet Oxygen Evolution In Battery Cathodes	83
6.1	Excited States of Molecular Oxygen	83
6.2	Marcus Theory of Electron Transfer	84
6.3	Solvent Effects on Reaction Energy	87
6.4	Understanding of Experimental Observation	92
6.4.1	Lithium Air Batteries	92
6.4.2	Sodium Air Batteries	93
6.4.3	Nickel Rich Lithium-Ion Batteries	94
6.5	Summary	97
7	Conclusion	98
8	Outlook	100
8.1	Transition Metal Cathodes	100
8.2	Behler-Parrinello Neural Networks	100
8.3	Beyond Fingerprint Based Learning	101

List of Tables

3.1	Table reproduced from Ref. [58] showing both the predicted and experimentally measured magnetic moment per magnetic ion, μ_B . Both the olivine (o) and quartz (q) polymorphs are included for LiFePO_4 . In the case of CuCr_2O_4 , with two magnetic species, the total magnetic moment is given. Experimental references are from (a) Ref.[69], (b) Ref. [70], (c) Ref. [71], (d) Ref. [72], and (e) Ref. [73]	27
3.2	Predictions for the magnetic ground state and magnetic energy difference $E_{FM} - E_{AFM}$ of various materials reproduced from Ref [58] the standard deviation, σ , and mean, μ , fro the BEEF-vdW ensemble is given along with the c-value incorporating all magnetic states, as well as the c-value derived from a Gaussian fit of the ensemble.]	30
3.3	The c-values for testing the hypothesis that the magnetic ordering is AFM over the selected length scale in the material. The last row for each material is the c-value testing the prediction of the most stable AFM ordering compared to all eight magnetic orderings tested. Also shown are the results from [83] for PBE and PBE+U with an onsite Hubbard correction of $U=0.95\text{eV}$ and $J=0.95\text{eV}$. The lengths given are the distances between Fe atoms in the FM structure and approximate the lengths between Fe atoms in the different AFM structures.	31
5.1	A comparison of training results on a subset of the NMC data from Chapter 4. The use of an output multi-layer perceptron (MLP) was tested versus liner regression on the learned feature of the convolutional layers.	82

6.1 Predicted ΔG^0 to form ground state O_2 for the free superoxide and cation associated superoxide reactions from Equations 6.4 and 6.5 respectively and the corresponding rate constant in various solvents: tetraethylene glycol dimethyl ether (TEGDME), dimethoxyethane (DME), dimethyl sulfoxide (DMSO), dimethylacetamide (DMA), dimethylformamide (DMF), and acetonitrile (MeCN). The solvent reorganization energy was calculated using Eq. 6.3 using $a=2.0 \text{ \AA}$ and $R=3.5 \text{ \AA}$ for the radius of each reactant and separation respectively. Data reproduced in part from Ref [177]. 90

List of Figures

2.1	Schematic of the artificial neural network to predict the energy of atom i using Behler-Parinello symmetry functions as input.	13
3.1	The total density of states ferromagnetically aligned LiFePO_4F as calculated by BEEF-vdW, PBE, RPBE, and PBEsol. [Reprinted with permission from Ref. [58]. Copyright 2017 American Physics Society]	24
3.2	The ensemble of energy difference predictions to determine the magnetic ground state of Cr is shown. The c -value for the prediction of AFM as the ground state is shown as the shaded region. Additionally, the normal distribution with ensemble mean and standard deviation is also shown and in good agreement with raw data. [Reprinted with permission from Ref. [58]. Copyright 2017 American Physics Society]	28
3.3	Schematic representing the procedure to calculate thermodynamic properties from the Debye model: First, an E-V curve (a) is fit and the free energies at different temperatures are calculated (b). The red line consists of the minimum of free energy at each temperature, from which the Gibbs energy (c), the volume (d), the bulk modulus (e), and vibrational entropy (f) are then computed. [Reproduced from Ref [59], with the permission of AIP Publishing, Copyright 2019]	33
3.4	Prediction of thermodynamic properties (a) Gibbs energy (b) entropy (c) isobaric heat capacity (d) bulk modulus (e) volume and (f) volumetric thermal expansion coefficient for Li using the Debye-Grüneisen Model compared with experimental measurements for Ref.[95, 96, 97, 98, 99]. For each property, the probability distribution from the BEEF-vdW ensemble is given. In (f), the natural log of the probability density is shown to demonstrate the extreme outliers. [Reproduced from Ref [59], with the permission of AIP Publishing, Copyright 2019.]	35

3.5	Prediction of thermodynamic properties (a) Gibbs energy (b) entropy (c) isobaric heat capacity (d) bulk modulus (e) volume and (f) volumetric thermal expansion coefficient for LiCoO_2 using the Debye-Grüneisen Model compared with experimental measurements for Ref. [100, 101, 102] cite, cite, cite. For each property, the probability distribution from the BEEF-vdW ensemble is given. In (f), the natural log of the probability density is shown to demonstrate the extreme outliers. . .	36
4.1	Schematic for a lithium-ion battery. On the left, we have the anode (denoted by -) which is conventionally graphite. On the right, we show the cathode (denoted by +) as a layered transition metal oxide. The lithium moves through the electrolyte from the anode to the cathode during discharge as the electron goes through the external circuit to provide usable electricity. During charge, an applied potential moves this process in reverse. The lithium moves out the cathode and back into the anode.	39
4.2	The O3 structure of $\text{LiNi}_{1-x-y}\text{Mn}_x\text{Co}_y\text{O}_2$ with rhombohedral symmetry given by the $R\bar{3}m$ space group. The transition metals are shown in green and form a triangular lattice. The oxygen in red form a cubic close-packed like structure. And the white show Li or lithium vacancies. These together also form a triangular lattice. The transition metals and the lithium sites are octahedrally coordinated with the oxygens leading to stacking of the layers such that they repeat every three layers, giving the O3 name.	41
4.3	A catalogue of conventionally used NMC (referred to as NCM here) compositions. Not the sparsity in the number of compositions used. [Reproduced from Argonne National Lab Copyright 2014.]	43
4.4	A cluster interactions considered up to 4 body within a cutoff of 6 Å.	46
4.5	Results for the four model selection techniques, (a) leave one out cross validation root mean squared error (RMSE), (b) the Akaike information criterion (AIC), (c) the the corrected Akaike information criterion (AICc), and (d) the Bayesian information critereon (BIC). [Reprinted from Ref [132], Copyright IOP Publishing 2020]	49
4.6	The prediction of energy as a function of composition for a one-dimensional slice of the total phase space. The ordered phases are given as blue dots and the disordered phases as a pink line. The large energy difference between the ordered and disordered phases demonstrates a large tendency for ordering. The multiple blue dots at a given composition are the energies of the ordered phases with different supercell sizes that were sampled. Reproduced from Ref [132], with the permission of IOP Publishing.	51

4.7	exp shown in red x from Ref. [135] disordered phases shown in dark red and ordered phases shown in blue. Computational data taken from Ref [132]. The high-Mn region determined to be unstable in the layered structure is blocked out.	52
4.8	The prediction of ordered and disordered phases and the associated confidence value (c-value) are shown in blue and dark red respectively. Again high-Mn region is blocked out and the experimental phases from Ref. [135] are shown. [Reproduced from Ref [132], with the permission of IOP Publishing Copyright 2020]	53
4.9	Radial Distribution Functions (rdf) for (a) NMC111 and (b) NMC811. For reference, the rdf for the all transition metals, which would represent a disordered phase in the supercells used are shown in (c) and (d). The comparison shows local ordering with less long ordering beyond 10 Å.	55
4.10	Predicted average voltage over all compositions of NMC. [Reproduced from Ref [132], with the permission of IOP Publishing, Copyright 2019]]	57
4.11	Dragonfly optimization. Reproduced from Ref [136], with the permission of AIP Publishing, Copyright 2019]	59
4.12	Plot of the force and energy RMSE on the test set as a function of number of epochs trained on the training set. Reproduced from Ref [136], with the permission of AIP Publishing, Copyright 2019]	61
4.13	Debye-Grüneisen model prediction from the BPNN trained in this chapter compared to the full DFT predictions from Figure 3.5 This thermodynamic predictions include (a) Gibbs energy (b) entropy (c) isobaric heat capacity (d) bulk modulus (e) volume and (f) volumetric thermal expansion coefficient for LiCoO ₂	62
4.14	Computed phonon spectrum for LiCoO ₂ using the BPNN calculator. .	64
4.15	Predicted open circuit voltages for a series of NMC compounds.[Reproduced from Ref [136], with the permission of AIP Publishing, Copyright 2019]	65
4.16	Prediction of open circuit voltage from the BPNN for (a) LiNiO ₂ compared with recent experiment from Ref, [110] and previous predictions using Cluster Expansion trained on DFT in Ref. [148], and (b) for LiCoO ₂ compared with experiment from Ref. [149] and previous predictions using Cluster Expansion trained on DFT in Ref. [124]. . . .	69

4.17	The various structures of LiNiO ₂ . The relation between the H1-1, O3, and O1 layer stackings are seen in (a). The whole picture shows the H1-3 pattern, while the dotted line shows the O1 stacking, and the solid line the O3 stacking. The correspondence between the hexagonal and monoclinic symmetries is seen in (b). And the phase diagram as measured from a collection of experiments as well as the predictions in blue from the computational phase modeling studying of Chang <i>et al.</i> The phase diagram for LiCo ₂ is largely similar. [Reprinted from Ref. [150]. Copyright 2012, with permission from Elsevier]	71
4.18	Test of the prediction of monoclinic distortions in LiNiO ₂ . For comparison, experimental predictions from Li <i>et al</i> [109] and cluster expansion trained on the a_m/b_m ratio by Arroyo y de Dompablo <i>et al.</i> [151]. We also show the line corresponding to a ratio of $\sqrt{3}$ which would indicate no distortion from the hexagonal phase.	72
4.19	Prediction of the lattice contractions that occur during the high states of charge for three NMC compositions all with 90% Ni content.	74
6.1	The log of the predicted rate of oxygen production in the singlet (dotted line) and triplet (solid line) spin state as a function of the change in Gibbs energy for the triplet ground state reaction with a reorganization energy of $\lambda = 1.3\text{eV}$. Note the inverted region in both curves, demonstrating that the rate does not continue to increase as the thermodynamic driving force increases. Thus the rate of singlet ultimately overtakes the rate of triplet oxygen production at a high driving force. [Reprinted with permission from [177]. Copyright 2020 American Chemical Society]	86
6.2	The linear fits for (a) the solvation energy of Li ⁺ as a function of donor number of the solvent, (b) the solvation energy of O ₂ ⁻ as a function of acceptor number of the solvent, (c) the solvation energy of LiO ₂ as a function of the sum of donor and acceptor number of the solvent, and (d) The dissociation energy of LiO ₂ into Li ⁺ and O ₂ ⁻ as a function of the combined solvation energies of the free ions. All numbers are for dimethoxyethane (DME), dimethylsulfoxide (DMSO), acetonitrile (MeCN), and dimethylacetamide (DMA) taken from Kwabi <i>et. al.</i> in Reference [182].	88
6.3	both (¹ Δ _g) and the (¹ Σ _g ⁺) spin singlet states considered. (a) $\lambda = 1.19$ eV to represent DME and (b) $\lambda = 1.62$ eV to represent MeCN. [Reprinted with permission from [177]. Copyright 2020 American Chemical Society]	91

6.4	Calculated reaction pathway (c) for the decomposition of Li_2CO_3 surface shown from the side (a) and from above (b). Lithium atoms are shown in purple, oxygen in red, and carbon in gray. [Reprinted with permission from [177]. Copyright 2020 American Chemical Society]	95
6.5	Schematic for the process of singlet oxygen production in Li-air and Li-ion batteries. Only the free, versus cation associated, superoxide disproportionation step is shown as that this key reaction in the generation of singlet oxygen. [Reprinted with permission from [177]. Copyright 2020 American Chemical Society]	96

Chapter 1

Introduction

The core of the thesis is a collection of methods for studying materials using first principles methods. These techniques could in principle be applied to a range of applications but are specifically applied to the applications of lithium-ion battery cathodes. The main goal of the methods used in the thesis is twofold. First is to provide a rigorous understanding of uncertainty at all levels, and best practices for model selection. And second is to increase the speed of computational predictions while increasing the error by an insignificant amount. To this end, a range of computer aided and machine learning techniques are employed, including shallow learning in the form of multivariate linear regression, and Bayesian statistics, and deep learning in the form of neural network potentials.

These two goals interplay throughout the work. Much of the work relies on calculations from density functional theory, which although well accepted has inherent uncertainty. Additionally, although DFT provides reasonable computational efficiency, when trying to search a large collection of materials as is done in binary or ternary phase search, or when trying to model large cells with hundreds of atoms, the cost of DFT could be prohibitive. To accelerate this, a more computationally efficient model can be trained at the cost of accuracy in the prediction with respect to the DFT. An understanding of the significance of the added error of this model on the predictions requires an understanding of the uncertainty of the DFT prediction itself.

In parallel, beyond their widespread use in consumer electronics, advances in lithium-ion battery technology show promise for the electrification of nearly all ground transportation. These electric vehicles, in conjunction with the constant increase in reliability and decrease in cost for renewable electricity, could enable the decarbonization of the transportation sector, a major contributor to global emissions. To reach a larger market share of electric vehicles, including long-haul trucking, an increase in capacity, cycle life, and energy density, accompanied by a decrease in cost is required. The most promising Li-ion cathode today is the layered form of Ni-Mn-Co (NMC) oxides with the formula $\text{LiNi}_x \text{Mn}_y \text{Co}_{1-x-y} \text{O}_2$. While dominating the market share of electric vehicles [1], recent concerns regarding the cost and availability

of Co [2] have pushed the materials towards higher Ni content. While bringing increased reversible capacity and lower cost over the high Co alternatives, a series of tradeoffs exist including decreased cycle life. Understanding trends in performance over the ternary phase space requires expensive experimentation. And the conventional performance based testing of cycling the battery over a set voltage range limits the understanding of what the material looks like at any given point during the experimentation. Structural changes accompany the changing lithium content which contributes to degradation. In LiNiO_2 as well as high-Ni NMC, the rapid change in the layer separation cause cracking in the materials and lost capacity [3]. And recent experiment has found at high states of charge the release of the highly reactive singlet spin state of molecular which is believed to evolve from the lattice oxygen of NMC [4]. The computational tools used here seek to further both the understanding and the speed at which new insight can be gained for these materials.

1.1 Overview of Thesis and Contributions

The goal of this thesis is to develop a set of tools based on first principles computation for the accurate and computationally efficient predictions and understanding of layered transition metal oxides in the context of lithium ion batteries. The purpose of these computational techniques is to understand the performance, structural changes, and degradation in layered transition metal cathode materials as a function of transition metal composition, and to enable uncertainty quantification in these predictions. We briefly preview the contributions of this thesis.

1.1.1 Uncertainty Quantification

Density functional theory is enabled vast progress in computational materials science, quantum chemistry, and solid state physics. Despite this success, a comprehensive and quantitative understanding of the uncertainty in the predictions remains a challenge. This is even more crucial when the predictions of density functional theory are passed through physics based models that may be highly sensitive to these errors. Within this thesis, a quantitative metric for prediction confidence based on an ensemble approach for the exchange correlation potential is introduced. This prediction confidence is discussed in the use case of the prediction of magnetic ground states.

The propagation of uncertainty through thermodynamic modeling is also explored. We employ the Debye-Grüneisen model for the prediction of the vibrational contributions to various thermodynamic properties, specifically the energy, at finite temperature. The ensemble approach used ensures proper treatment of the statistical dependence of the correlated errors of density functional theory. From our results, we discuss the sensitivity of the derived properties to the underlying error in the energy inputs of the model.

1.1.2 Voltage Prediction

In the context of batteries, the most fundamental and useful prediction that can be made for a battery is that of the open circuit voltage. This represents a prediction of how much energy can be extracted, how consistent the voltage is over the operating range, and how much capacity can be extracted when cycling to certain cutoff voltages. In real operation, the range over which the battery is cycled controls the long term performance of the cathode. A theoretical mapping of battery voltage to lithium amount of lithium in the battery can enable insight into the mechanism for worse cycling performance when using higher cutoff voltages. Previous predictions of this voltage curve have been done on a case by case basis with models that are specialized to only a single cathode composition. Therefore the systematic comparison across many cathode compositions was previously not possible. In this thesis, we present a method for the generation of these open circuit voltages for any NMC cathode composition using Monte Carlo simulations and an advanced neural network potential for the accurate prediction of energy and forces.

1.1.3 Battery Degradation

Beyond the ideal open circuit voltage, the prediction of structural changes can aid in the insight of degradation mechanisms that occur at high states of charge in these cathode materials. Most specifically, the rapid lattice contraction for high-Ni NMC leads to cracking and increases surface reactivity. The accurate machine learning potential can accurately predict this lattice contraction as well as explore the effect of Co and Mn compositions in trying to mitigate it.

A parallel degradation mechanism that also occurs at when charging these cathodes to high potentials is the release of highly reactive molecular oxygen in the excited singlet spin state. The decay to the spin triplet ground state is forbidden by spin transition rules and therefore will instead react with and ultimately degrade the battery electrolyte. The search for mechanism for the appearance of this oxygen become more intriguing with the appearance in Li-air and Na-air battery technologies. Within this thesis, we elucidate this mechanism for singlet oxygen production, therefore unifying a degradation mechanism shared by these three important battery technologies. Additionally, our analysis is able to rationalize all experiments on the observation of singlet within these fields.

1.2 Organization of Thesis

The organization of this thesis is as follows. We begin in Chapter 2 with a discussion of the methods used within this work. We start by discussing the theory of DFT and various approaches to the largest approximation, the exchange correlation functional. We then discuss the mathematical framework of featurizing materials using atom

centered symmetry functions to train a neural network on the prediction of energy and forces. We close with a derivation of the Debye model as well as the corrections used here for more accurate predictions of thermodynamic properties.

Chapter 3 dives into the uncertainty of density functional theory. We utilize the Bayesian Estimation Functional with van der Waals correction (BEEF-vdw) to understand the confidence in ground state predictions over various configurations of a material and to understand the sensitivity of DFT error in the predictions of the Debye model.

In Chapter 4, we introduce the material space of NMC cathodes and discuss why Ni, Mn, and Co have emerged as the choice of transition metals to be used. The properties of this material space are explored first with a search of the transition metal ordering followed by the prediction of open circuit voltage and lattice variations during cycling with the use of a machine learning potential discussed in this chapter. We also test the predictions of thermodynamic properties from this machine learning potential and related it to the results of Chapter 3.

In Chapter 5, we quickly discuss the field of machine learning potentials with a review of the many current methods used today. We discuss in more detail covariant convolutional neural networks and the extension of a molecular network architecture of this kind to the prediction of energy and forces on periodic systems.

The mystery of singlet oxygen production is explored in Chapter 6 where with the use of theory for kinetics, we present an understanding of the mechanism and influence of electrolyte on the production of singlet oxygen in Li-air, Na-air, and NMC batteries. Insight is then drawn from this analysis to explain experimental observations.

We close with a conclusion of the work in Chapter 7 and an outlook in Chapter 8. We discuss the current state of Ni-rich NMC as well as the possibilities and limitations for commercialization. We also present a discussion of the many opportunities in computational battery research the machine learning potential presented in this thesis can provide. And finally, we look to the future of machine learning potentials with a discussion of the possibilities advanced architectures present and insight into data generation for training depending on the desired use of the potential.

Chapter 2

Methods

This chapter describes the most integral methods used in this thesis. We first begin with a discussion of Density Functional Theory (DFT), a well tested computational technique to predict the properties of material given only the atomic identities and positions. We briefly move through the reformulation of the intractable full many-body problem to the more easily applied set of Kohn-Sham equations which can be solved self consistently for the electron density and energy. We also discuss the most important approximation of Density Functional Theory, the exchange-correlation potential, and some of the most successful and most used approximations to this functional. Next, is a discussion of the application of machine learning methods which can be trained on the outputs of DFT in order to provide a surrogate model that is an order of magnitudes faster, with negligible loss of accuracy. We focus on a featurizing to the input materials using a set of atom centered Gaussian symmetry functions suggested by Behler and Parinello. We end the chapter with a discussion of the Debye-Grüneisen Model for the prediction of the vibrational properties of materials. This approximation of the phonons within a material is able to generate accurate results with respect to experiment as we will see later and only requires a set of energy calculations over various volumes. We first present the original Debye model for elastic medium and then move on the discuss corrections to this model including a scaling factor to account for the variation in transverse and longitudinal velocities, and the use and approximations of the Grüneisen parameter.

2.1 Density Functional Theory

The energetics and dynamics of atomic nuclei and their accompanying electrons can be well described by the non-relativistic Schrodinger equation.

$$\hat{H}\Psi = i\frac{\partial}{\partial t}\Psi \tag{2.1}$$

where Ψ is the many-body wavefunction and \hat{H} is the time-dependent Hamiltonian

of the system containing M nuclei and N electrons and is given by

$$\hat{H} = -\sum_{i=1}^N \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_{A=1}^M \frac{\hbar^2}{2M_A} \nabla_A^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{1}{4\pi\epsilon_0} \frac{q_e^2 Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{4\pi\epsilon_0} \frac{q_e^2}{r_{ij}} + \sum_{A=1}^M \sum_{B>A}^M \frac{1}{4\pi\epsilon_0} \frac{q_e^2 Z_A Z_B}{R_{AB}}. \quad (2.2)$$

Here the indices i and j run over all electrons in the system and A and B denote all nuclei in the system. The first two terms of this Hamiltonian describe the kinetic energy of the electrons and the nuclei, while the remaining terms are the electrostatic interactions of the electrons and nuclei, the electrons with other electrons, and nuclei with other nuclei. Given the relatively large difference in mass between the electrons and nuclei, the nuclei will move much slower than the electron and the Born-Oppenheimer approximation can be made wherein the positions of the nuclei are assumed stationary. The properties of a material or molecule can then be understood as a series of electronic structure solutions for fixed ionic positions by solving the electron only time-independent Schrodinger equation $H_e \Psi = E_e \Psi$. The Hamiltonian is then reduced to

$$\hat{H}_e = -\sum_{i=1}^N \frac{\hbar^2}{2m_e} \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{1}{4\pi\epsilon_0} \frac{q_e^2 Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{4\pi\epsilon_0} \frac{q_e^2}{r_{ij}}. \quad (2.3)$$

The total energy of the system is now the sum $E_{tot} = E_e + E_n$ where the nuclear energy E_n is given by the electrostatic nuclear interaction energy. Additionally, as this Hamiltonian is uniquely defined for the external potential generated by the ions, the ground state of the system must also be uniquely determined by this external potential. In practice, solving for this ground state and the corresponding electronic wavefunctions of a many-electron system in an external potential is intractable and an alternative approach is needed.

2.1.1 Kohn-Sham Equations

In 1964 Hohenberg and Kohn were able to successfully reduce the dimensionality of the problem from 3N dimensions to just 3 dimensions by mapping the problem of finding a many-body wavefunction to that of solving for the electron density of the interacting electron system.[5] They showed that the external potential v_{ext} is determined uniquely up to a constant by the ground state density of the electrons $n_0(\mathbf{r})$. Thus, since the external potential determines the Hamiltonian and the ground state, all physical observables, most importantly the energy, of the ground state can in principle be determined given the ground state electronic density. They also proved that there exists a universal functional of the energy $E[n]$ that is valid for any

external potential, v_{ext} . Therefore, for a given external potential, the ground state of the system is determined by the absolute minimum of this functional minimizing over all wavefunctions with density $n(\mathbf{r})$:

$$E_0[n] = \min_{\Psi \rightarrow n(\mathbf{r})} \langle \Psi | \hat{H} | \Psi \rangle = \min_{n(\mathbf{r})} E[n(\mathbf{r})] \quad (2.4)$$

This means that all that is needed to understand the ground state electronic structure of any atoms, molecule, crystal, or material is the identity and positions of the ions in that material in order to calculate the external potential felt by the atoms. This universal functional can be broken into two terms to separate parts, one that is universal to all systems and one that depends on the external potential.

$$E[n(\mathbf{r})] = F[n(\mathbf{r})] + \int v_{ext}(r)n(\mathbf{r})d\mathbf{r} \quad (2.5)$$

This universal functional, $F[n(\mathbf{r})]$ is not explicitly known but is defined as the minimum of the kinetic energy and electron-electron interaction energy over all wavefunctions of density $n(\mathbf{r})$:

$$F[n(\mathbf{r})] = \min_{\Psi \rightarrow n(\mathbf{r})} \langle \Psi | \hat{T} + \hat{V}_{ee} | \Psi \rangle \quad (2.6)$$

While this gives an exact description of the system, and seemingly the problem of minimizing over a 3N-dimensional wavefunction Ψ has been transformed to the much easier minimization over 3-dimensional electron density, the problem of finding the universal functional $F[n(\mathbf{r})]$ itself requires solving the 3N-dimensional problem to be avoided. In most real implementations of density functional theory, the Kohn-Sham approach is used by replacing the original many-body problem with a fictitious system of non-interacting electrons in an effective potential. This effective potential must then be constructed in such a way so that it ensures the fictitious non-interacting and the real interacting system have the same ground-state density.

In an approach similar to this, Hartree proposed in 1928 a set of self-consistent single-particle equations in an attempt to approximate a system of interacting electrons [6]. This was done by assuming that every electron was moving in an effective single-particle potential given by interactions with a nucleus of atomic number Z and the potential due to the average electron density $n(\mathbf{r})$

$$v_H = -\frac{1}{4\pi\epsilon_0} \frac{q_e Z}{|\mathbf{r}|} + \int \frac{1}{4\pi\epsilon_0} \frac{q_e n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}' \quad (2.7)$$

The ground state of the system can then be found by self-consistently solving the set of single-particle Schrodinger-like equations that are obtained from the variational principle when trying to minimize Hartree's approximation to the energy:

$$\left\{ \frac{\hbar^2}{2m} \nabla^2 + v_H(\mathbf{r}) \right\} \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}) \quad (2.8)$$

$$n(\mathbf{r}) = \sum_i^N |\psi_i(\mathbf{r})|^2 \quad (2.9)$$

While this approximation works well, it is unfortunately not exact. Drawing inspiration from this, we return to the Kohn-Sham formalization of the problem in an attempt to recreate the exactness of the Hohenberg and Kohn's original work, we can then rewrite Eq. 2.6 in terms of the non-interacting system.

$$F[n(\mathbf{r})] \equiv T_0[n(\mathbf{r})] + \frac{1}{2} \int \frac{1}{4\pi\epsilon_0} \frac{q_e^2 n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + E_{xc} \quad (2.10)$$

where the first two terms in this equation are the kinetic energy of the non-interacting system and the Hartree potential. The last term is the so called exchange-correlation energy which is defined in this equation as the remaining quantum mechanical energy difference between the exact functional $F[n(\mathbf{r})]$ of the interacting system and the known Hartree-like treatment of the fictitious non-interacting system. The value of the exchange correlation-energy functional is not known exactly and the accuracy of the density functional method thus depends largely on the approximation used for this term. The full energy to now minimize is given by

$$E = T_0[n(\mathbf{r})] + \int v_{ext}(\mathbf{r})n(\mathbf{r})d\mathbf{r} + \frac{1}{2} \int \int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + E_{xc}[n(\mathbf{r})] \quad (2.11)$$

Rewriting the exchange-correlation energy as $E_{xc}[n(\mathbf{r})] = \int n(\mathbf{r})\epsilon_{xc}d\mathbf{r}$, the variational principle leads to the Kohn-Sham Equations[7] that are solved self consistently:

$$\left\{ \frac{1}{2}\nabla^2 + v_{ext}(\mathbf{r}) + \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta\epsilon_{xc}}{\delta n} \right\} \psi_i(\mathbf{r}) = \epsilon_i \psi_i(\mathbf{r}) \quad (2.12)$$

$$n(\mathbf{r}) = \sum_i^N |\psi_i(\mathbf{r})|^2 \quad E = \sum_i^N \epsilon_i \quad (2.13)$$

This can be easily expanded to spin-polarized calculations by solving for the spin up and spin down Kohn-Sham wavefunctions separately and simultaneously.

$$\left\{ \frac{1}{2}\nabla^2 + V(\mathbf{r}) + \sum_\alpha \int \frac{n_{\alpha\alpha}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} + \frac{\delta\epsilon_{xc}}{\delta n} \right\} \begin{pmatrix} \phi_i^{(+)}(\mathbf{r}) \\ \phi_i^{(-)}(\mathbf{r}) \end{pmatrix} = \epsilon_i \begin{pmatrix} \phi_i^{(+)}(\mathbf{r}) \\ \phi_i^{(-)}(\mathbf{r}) \end{pmatrix} \quad (2.14)$$

$$n_{\alpha\beta}(\mathbf{r}) = \sum_i^N (\mathbf{r}) \phi_i^{*\alpha} \phi_i^\beta(\mathbf{r}) \quad (2.15)$$

where α , and β can be + or -. If the system is assumed to have collinear spin, the magnetization is given by [8]

$$M_{spin} = \int (n^{(+)}(\mathbf{r}) - n^{(-)}(\mathbf{r}))d\mathbf{r} \quad (2.16)$$

2.1.2 Exchange Correlation Functionals

The discussion so far has focused on reformulating the unsolvable many-electron problem into a seemingly more approachable formulation in terms of electron density. For stationary solutions to the full non-relativistic Schrodinger equation, this Kohn-Sham formalism has remained exact within the Born-Oppenheimer approximation. This is done, however, only by defining the exchange-correlation energy to be the universal function which includes all extra physics not already included, and that properly corrects the kinetic energy from that of the non-interacting fictitious system to that of the real many-electron system to be modeled. A series of approximations to this exchange-correlation functional have enabled predictions of reasonable accuracy at computationally viable costs.

In the original work of Kohn and Sham, this exchange-correlation functional was assumed to have a quasi-local form. This approximation, which is now known as the local density approximation (LDA), estimates the exchange-correlation at every point to be that of a homogeneous electron gas (HEG) of the same density. Given ϵ_{xc}^{HEG} is the exchange-correlation energy per electron in a HEG of density n ,

$$E_{xc}^{LDA}[n(\mathbf{r})] = \int n(\mathbf{r})\epsilon_{xc}^{HEG}[n(\mathbf{r})]d\mathbf{r} \quad (2.17)$$

While $\epsilon_{xc}^{HEG}[n(\mathbf{r})]$ is unique, several parameterizations of LDA have been proposed.[9, 10, 11, 12] This first approximation, despite its simplicity, has had remarkable success in many systems. It is able to predict ionization energies of atoms, dissociation energies of molecules, and cohesive energy of solids in many cases within 10% error. Additionally, it can predict geometries within 1% error.[13] It is especially accurate for elemental metals, as the nearly-free electrons within these materials largely resemble a slowly varying electron gas, where LDA is most applicable.

A more accurate estimation of the exchange-correlation energy depends on the gradient of electron density and is known as the generalized gradient approximation (GGA). The exchange-correlation energy is now approximated as

$$E_x^{GGA}(n, \nabla n) = \int \epsilon_{xc}[n, \nabla n]d\mathbf{r} \quad (2.18)$$

The freedom in the mathematical implementation of the gradient dependence has led to many versions of GGA with some notable versions including PBE[14], RPBE[15], and PW91 [16]. While GGA functionals have improved performance in predicting the properties of materials with more localized electrons including transition metal complexes, surfaces, and interfaces, the approximation for the exchange-correlation functional still assumes a quasi-local form. Further improvement of the energy and structure predictions can be made when adding an explicit non-local van der Waals density function (vdW-DF) kernel for the correlation energy.

$$E_c^{nl} = \frac{1}{2} \int r' n(\mathbf{r}) \phi(\mathbf{r}, \mathbf{r}') n(\mathbf{r}) d\mathbf{r} d\mathbf{r}' \quad (2.19)$$

The full correlation energy is now given by

$$E_c[n] = E_c^0[n] + E_c^{nl}[n] \quad (2.20)$$

Several non-local kernels have been developed with various success including DF [17] DF2 [18] vdW-DF-cx [19], and optPBE-vdW[20]. This addition of non-local correlation is important trying to accurately describe the interactions between layers of the layered transition metal oxide materials within this thesis [21].

2.2 Machine Learning Potentials

The tremendous success density functional theory has enabled computational predictions in high-throughput studies with the potential of accelerating material discovery for a wide range of applications, including in energy materials [22], thermoelectrics [23], electrocatalysts [24], topological insulators [25], magnetic materials [26], and solar materials [27]. Unfortunately, the computational requirements of DFT limit the size of and/or the number of materials that can be reasonably screened. To overcome this, advancements in machine learning (ML) techniques can be leveraged to create surrogate models for the predictions generated with density functional theory in the form of a machine learning potentials. While machine learning has been wildly successful in its ability to learn complex patterns from datasets, the issue of interpretability remains an issue in its application to the physical sciences. Its application to learning the DFT mapping of atomic structure to energy, however, may not require as much interpretability. The use of DFT is so widespread that in many cases, it is used as a black box with no easily extractable understanding can be derived explicitly from the inner workings. That is to say in replacing DFT with a computationally cheaper calculator of the same predictions of DFT, one loses very little in interpreting how DFT got to that answer and in some cases of model choice, may even gain understand (such as understanding the magnetic coupling of spins). This section will be dedicated to a summary of the theory and methods related to the type of neural network machine learning potentials used within this thesis.

2.2.1 Atom Centered Symmetry Functions

We focus on atom centered symmetry functions fed to artificial neural networks. These atom centered symmetry functions are used as they present a good trade-off between computational efficiency and accuracy [28] and have previously been shown to accurately describe many-component layered transition metal oxide systems [29]. There has been widespread success in using feed-forward neural networks to represent potential energy surfaces (PES) of various materials. For low dimensional potential energy surfaces, a naive implementation directly from atomic coordinates or explicit degrees of freedom can be achieved [30]. As the number of degrees of freedom increases, this becomes computationally prohibitively expensive as each degree

of freedom represents an input node to the neural network. Additionally, the neural network size and architecture cannot be changed once trained and therefore has to be retrained if the system size of the material to be predicted changes. To overcome this, it is useful to decompose the energy into a sum of contributions from each atom and train a series of neural networks for each atom’s energy contribution.

$$E(\mathbf{r}) = \sum_i E_i(\mathbf{r}_i) \quad (2.21)$$

Several physics requirements must also be satisfied for training the neural networks to represent the energy. The energy of the input material is invariant to rotations and translations within the coordinate system used to describe the positions of the atoms. Additionally, the energy is also invariant to the labeling of the atoms. The model for the potential energy surface should therefore also be invariant to these rotations, translations, and permutations. To do this, a set of functions are used to map the coordinates of the atoms to a description of the chemical environment that is invariant to these symmetries. The descriptors used in this thesis are the atom centered symmetry functions suggested by Behler and Parrinello [31, 32] given by

$$G_i^2 = \sum_{j \neq i} e^{-\eta(r_{ij}-r_s)^2/r_c^2} f_c(r_{ij}) \quad (2.22)$$

and

$$G_i^4 = 2^{1-\zeta} \sum_{j \neq i} \sum_{k \neq j} (1 + \lambda \cos \theta_{ijk})^\zeta e^{-\eta(r_{ij}^2+r_{jk}^2+r_{ik}^2)/r_c^2} \times f_c(r_{ij})f_c(r_{jk})f_c(r_{ik}) \quad (2.23)$$

where

$$f(r_{ij}) = \begin{cases} 0.5 \left[\cos \left(\frac{\pi r_{ij}}{r_c} \right) + 1 \right] & r_{ij} \leq r_c \\ 0 & \text{otherwise} \end{cases}$$

The two-body interactions of an atom i with all other atoms j within a cutoff distance, r_c is captured by the first symmetry function used, G_i^2 , while the G_i^4 symmetry function captures the three-body interactions of i with all j and k such that all of the atomic separations r_{ij} , r_{ik} , and r_{jk} are less than the cutoff distance and where θ_{ijk} is the angle formed by r_{ij} and r_{ik} . The parameters η and ζ are used to tune the distance and angle at which the symmetry function is largest and therefore will tune the focus of the symmetry functions to various interaction lengths and angles. The final tunable parameter, λ has a typical value of ∓ 1 as these values correspond to a maximum in the symmetry function at 0° for $\lambda = 1$ and 180° for $\lambda = -1$. To create a more robust description of the material space of interest, a collection of these symmetry functions with various η , ζ , and λ can be used in combination as inputs to generate a robust feature vector for each energy neural net. Within this thesis,

the values of $\lambda = \pm 1$ and $\zeta=4$ will be used as these values are shown to describe the angular distribution well. We can then use a collection of η values to represent the spacial component of interaction in the system.

The symmetry functions are then fed to a neural network (NN) consisting of several layers. Each layer and then comprised of several nodes, each with a numerical value. The input layer is given by the numerical values of each symmetry function used as the description of the atomic environment of the material. The node values, y_i^j of node i in layer j after the input layer is then computed as:

$$y_i^j = f_i^j \left(b_i^j + \sum_k a_{ki}^{j-1,j} \cdot y_k^{j-1} \right) \quad (2.24)$$

where k is summed over all nodes from the previous layer, y_k^{j-1} is the value of node k of the previous layer, $a_{ki}^{j-1,j}$ is a learnable weight between node k of the previous layer and the current node i of the current layer j , and b_i^j is learnable bias for layer j . Finally, f_i^j is an activation function that introduces a nonlinearity to the neural network and within this thesis, a sigmoid is used as the activation function

$$f(x) = \frac{1}{1 + e^{-x}}. \quad (2.25)$$

A schematic of the neural network architecture can be seen in Fig 2.1 to demonstrate the prediction of the energy from atom i in a single material with N atoms, and M descriptor functions. The full model for this material's energy is then a collection of neural networks summed up. As mentioned previously, the prediction of the final energy must be invariant to the labeling of atoms, therefore, instead of training a neural network for each atom in the system, one NN is trained for every atomic species in the system.

The learnable weights and biases are trained to minimize the loss function given by the sum of squared errors between the predicted energy \hat{E}_j and the target energy E_j :

$$\mathcal{L} = \frac{1}{2M} \sum_{j=1}^M \left(\frac{\hat{E}_j}{N_j} - \frac{E_j}{N_j} \right)^2. \quad (2.26)$$

The final component of a machine learning potentials is the prediction of forces as they are necessary to accurately describe the positions and distortion of atoms under various conditions. These force on each atom can be easily calculated by taking the derivative of the predicted energy with respect to the position of the atom. As the descriptors are analytical functions of the atomic positions, this can be done easily

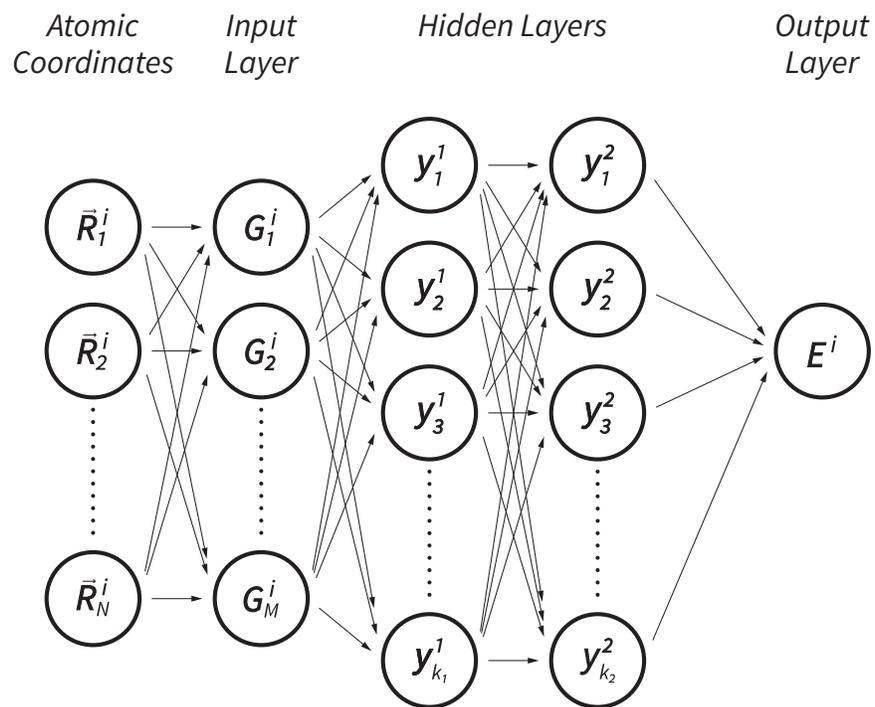


Figure 2.1: Schematic of the artificial neural network to predict the energy of atom i using Behler-Parinello symmetry functions as input.

using the chain rule:

$$F_{k,\alpha} = -\frac{\partial E}{\partial R_{k,\alpha}} = -\sum_{i=1}^N \frac{\partial E_i}{\partial R_{k,\alpha}} \quad (2.27)$$

$$= -\sum_{i=1}^N \sum_{j=1}^{M_i} \frac{\partial E_i}{\partial G_{i,j}} \frac{\partial G_{i,j}}{\partial R_{k,\alpha}}. \quad (2.28)$$

Finally, force training requires modification of the loss function to balance the weighting of the error in the prediction of energy with the error in the prediction of forces. We, therefore, redefine the loss function used to train the neural network as

$$\mathcal{L} = \frac{1}{2M} \sum_{j=1}^M \left[\left(\frac{\hat{E}_j}{N_j} - \frac{E_j}{N_j} \right)^2 + \frac{\alpha}{3N_j} \sum_{k=1}^3 \sum_{in=1}^{N_j} (\hat{F}_{ik} - F_{ik})^2 \right], \quad (2.29)$$

where \hat{E}_j and E_j are the predicted and reference energy of the j^{th} data point, \hat{F}_{ik} and F_{ik} are the predicted and reference force on atom i of the j^{th} data point in the k direction, and α is the force coefficient that sets the relative weighting between the force and energy sum of squared errors. Unfortunately in the application of training a machine learning potential based on these atom center symmetry functions on both energy and forces, there is an inherent trade-off. Increasing the force coefficient will increase the accuracy of the potential in recreating forces at the cost of a decrease in energy prediction accuracy.

2.3 Debye-Grüneisen Model

Density functional theory can be useful for the prediction of ground state zero temperature energy of materials. In many cases, however, finite temperature effects due to vibrations of the lattice play a large role in the energy, phase stability, volume expansion, heat capacity, and more of a material. Most methods to calculate the vibrational properties such as with finite-difference [33], or with density functional perturbation theory [34] are too computationally expensive for high-throughput calculations and a more approachable approximation is needed. The Debye-Grüneisen model provides a computationally efficient way to compute approximate these vibrational effects and has been applied to a large range of materials including oxides, metals, nitrides, carbides, fluorides, and sulfides [35, 36, 37, 38]

2.3.1 Original Debye Model

The vibrational properties of materials can be understood through the collective vibration of atoms within a crystal in the form of quantized excitations known as

phonons. The Debye model assumes a linear dispersion relation for these phonons $\omega = v_s |\vec{q}|$ where v_s is the speed of sound in the material.

As these phonons are merely vibrations within the material, they cannot have infinite frequency and are bound by the size of the lattice, and they additionally cannot have infinite energy $E = \hbar\omega$. A cutoff frequency called the Debye frequency ω_D and cutoff wavevector $q_d = \omega_d/v_s$ can be used to approximate the limits of these vibrations. The values of these cutoffs can be computed by approximating the material as a cubic crystal of size length $L = (V)^{1/3}$. Therefore the wave-vectors take the form of

$$q = \frac{2\pi n}{L} \quad (2.30)$$

where n is an positive, non-zero integer. Therefore in reciprocal space, each wavevector \vec{q} has a volume of

$$\left(\frac{2\pi}{L}\right)^3 = \left(\frac{8\pi^3}{V}\right) \quad (2.31)$$

Given the cutoff wavevector q_D , the total number of modes N , where n is the number of primitive cells in the material, confined within a sphere of radius q_D [39].

$$N = \left(\frac{4}{3}\pi q_D^3\right) / \left(\frac{8\pi^3}{V}\right) \quad (2.32)$$

This then leads to the value of the Debye frequency

$$\omega_D = v_s \left(6\pi^2 \frac{N}{V}\right)^{1/3} \quad (2.33)$$

It is also helpful to define the Debye temperature, Θ_D , above which, all of all modes will begin to be excited,

$$k_B \Theta_D = \hbar\omega_D \quad (2.34)$$

Combining Equations 2.33 and 2.34 and restricting ourselves to only one primitive cell, we then have

$$\Theta_D = \left(\frac{6\pi^2}{V}\right)^{1/3} \frac{\hbar v_s}{k_B} \quad (2.35)$$

The density of states for each polarization type is then given as

$$g(w) = \frac{dN}{d\omega} = \frac{V\omega^2}{2\pi^2 v_s^3} \quad (2.36)$$

Assuming a constant velocity for all directions, we multiply the contribution for a single polarization by three to get the total contribution to the internal energy from vibrations[39]

$$U = 3 \int \hbar\omega_D(\omega)g(\omega) \left(n(\omega) + \frac{1}{2} \right) d\omega = 3 \int_0^{\omega_D} \frac{V\omega^2}{2\pi^2v_s^3} \left[\frac{\hbar\omega}{e^{\hbar\omega/k_B T} - 1} + \frac{\hbar\omega}{2} \right] d\omega \quad (2.37)$$

where we have made the substitution for the expected number of particles from Bose-Einstein statistics $n\omega$. We can also break the energy into the contribution from the second term, the zero point-energy, U_0 , and the first term which is the temperature-dependent vibrational energy. Making the substitution $x = \hbar\omega/k_B T$ and $x_D \equiv \Theta_D/T$ the temperature-dependent energy can be written

$$U - U_0 = 9Nk_B T \left(\frac{T}{\Theta_D} \right)^3 \int \frac{x^3}{e^x - 1} dx = 3k_B T D(\Theta_D/T) \quad (2.38)$$

Where for notational convenience, we have used the Depye function $D(x) = \frac{3}{x^3} \int_0^x \frac{z^3 dz}{e^z - 1}$. The zero point energy can be represented as

$$U_0 = \frac{9}{8} k_B \theta_D. \quad (2.39)$$

From this internal energy, we have a full thermodynamic potential and can then compute all relevant quantities of interest. We start by computing the isochoric heat capacity as

$$C_V = \frac{\partial U}{\partial T} = 9Nk_B \left(\frac{T}{T_D} \right) \int_0^{x_D} \frac{x^4 e^x}{(e^x - 1)}. \quad (2.40)$$

To ultimately reach an expression for the free energy, we need to next compute the entropy which is given by the relation $C_v = T \frac{\partial S}{\partial T}$ and thus

$$S = \int_0^T \frac{C_V}{T'} dT' = 3k_B \left[\frac{4}{3} D(\Theta_D/T) - \ln(1 - e^{-\Theta_D/T}) \right]. \quad (2.41)$$

The last thermodynamic potential to be derived here will be the vibrational Helmholtz free energy as this will be the starting point for the application of the Debye model later. The expression for this is given by

$$F = U - TS = -k_B T \left[D(\Theta_D/T) - 3\ln(1 - e^{-\Theta_D/T}) \right] + \frac{9}{8} k_B \theta_D. \quad (2.42)$$

The Helmholtz free energy can be numerically calculated for any temperature and any volume and added to the zero temperature static energy at that volume calculated from DFT. Before this can be done, the tabulation of the Debye temperature requires an approximation for the speed of sound in the material. For the moment, we can

assume that the excitations in all directions within the material have a constant speed of sound which can be given by the Bulk modulus and density as

$$v_s = \sqrt{B/\rho}. \quad (2.43)$$

To then give

$$\Theta_D = (6\pi^2)^{1/3} V^{1/6} \frac{\hbar}{k_B} \left(\frac{B}{M} \right)^{1/2} \quad (2.44)$$

The Debye Temperature can now be calculated easily from first-principles calculations as it is in the form of easily computable quantities, the equilibrium volume V , and the bulk modulus B . These two quantities are easily extracted from utilizing the energy-volume equation of state curves fit to DFT data. The equilibrium volume is simply given from the energy minimum in the energy-volume curve. The bulk modulus on the other hand is given by the curvature of the energy-volume fit. The bulk modulus is defined in terms of the volume change as the pressure is varied

$$B = -V \frac{\partial P}{\partial V} \quad (2.45)$$

We rewrite the pressure,

$$P = -\frac{\partial E}{\partial V} \quad (2.46)$$

and insert this into the original definition of bulk modulus to get

$$B = V \frac{\partial^2 E}{\partial V^2}. \quad (2.47)$$

This simple formalism thus only requires enough single point energy calculations required to give a sufficient energy-volume curve. From there, all possible thermodynamic quantities can be computed. We apply this method first in Chapter 3 to understand the sensitivity of the DFT error on the prediction results, and then later to cathode materials in Chapter 4. The approach taken, which ultimately seeks to predict the Gibbs Free energy and is described in more detail in Chapter 3, is to compute the Helmholtz free energy at a collection of temperatures and then compute the Gibbs energy from the Legendre transform $G = F + PV$.

2.3.2 Anharmonic Corrections

We will now discuss some corrections to the expression of the Debye temperature that provides an even more accurate and flexible description of materials. Firstly, the approximation for the speed of sound used leads to Debye temperatures much larger than experimentally measured. [35] This can be corrected by multiplying by an empirical scaling factor that accounts for the difference in the speed of sound in the

transverse and longitudinal directions that are present even for an isotropic medium. The average speed in the material can be expressed in an averaged way in terms of the longitudinal and transverse speeds as[40]

$$v = \left[\frac{1}{3} \left[\frac{2}{v_t^3} \right] + \left[\frac{1}{v_l^3} \right] \right]^{1/3} \quad (2.48)$$

where $v_t = \sqrt{S/\rho}$ and $v_l = \sqrt{Y/\rho}$. Here S and Y are the shear and longitudinal (Young's) moduli respectively. In an isotropic medium, these two moduli can be written in terms of the Poisson ratio, ν , and the bulk modulus, B , as [41]

$$Y = \frac{3(1-\nu)}{1+\nu} B \quad (2.49)$$

$$S = \frac{3(1-2\nu)}{2(1+\nu)} B \quad (2.50)$$

We now obtain a new form of Equation 2.43 multiplied by a scaling factor, s , that modifies the approximated speed of sound in the materials $v = s(\nu) \sqrt{\frac{B}{\rho}}$ where

$$s(\nu) = \left(\frac{1}{3} \left[\left(\frac{1+\nu}{3(1-\nu)} \right)^{3/2} + 2 \left(\frac{2(1+\nu)}{3(1-2\nu)} \right)^{3/2} \right] \right)^{-1/3}. \quad (2.51)$$

When the Poisson ration is not known a value of $s=0.617$ has been suggested [35]. Anharmonic effects in the vibrations of the lattice can then be accounted for in the form of the Grüneisen parameter given by[35]

$$\gamma = - \frac{\partial \ln \Theta_D}{\partial \ln V} \quad (2.52)$$

Put simply, this parameter describes the effect of changing the lattice volume has on the vibrational properties of the material. The value of the Grüneisen parameter can found by plugging in the expression for the Debye temperature in Equation 2.44 to yield

$$\gamma = -\frac{1}{6} - \frac{1}{2} \frac{\partial \ln B}{\partial \ln V} = -\frac{1}{6} - \frac{1}{2} B' \quad (2.53)$$

where B' is the dimensionless derivative of the bulk modulus with respect to the pressure. This expression assumes that both the transverse and longitudinal modes will be equally excited and each type of mode will vary in the same way with respect to volume. This is of course valid at high temperature, above the Debye temperature, where all of the modes are excited. At low temperatures, this expression for γ has been shown to be too large by an additive factor of about $\frac{1}{3}$ [42]. Therefore, the lower temperature Grüneisen parameter can be approximated as

$$\gamma = -\frac{1}{3} - \frac{1}{2}B' \quad (2.54)$$

The volume dependence finally enters into the Debye temperature through the relation [35] $\Theta_D/\Theta_D^0 = (V/V_0)^\gamma$ and the final expression for the Debye temperature is

$$\Theta_D = s(6\pi^2)^{1/3}V_0^{1/6} \frac{\hbar}{k_B} \left(\frac{B}{M}\right)^{1/2} \left(\frac{V_0}{V}\right)^\gamma \quad (2.55)$$

In summary, we have now have an approximation for the Debye temperature that takes into account the variation in transverse and longitudinal sound speeds. It also accounts for the changing vibrational properties as the volume of the material changes from the zero temperature volume, as will happen for most material when going to finite temperatures of interest. The approximation still relies on the assumption of isotropic interaction in the material. Therefore if a material varies largely in different crystallographic directions, or has significant contributions to the vibrational properties that derive to two distinct bonding types (such as molecular crystals that have strong covalent bonding at short range, and weaker van der Waals interaction in the intermolecular binding. Additionally, corrections to the free energy from electronic entropy could play a role in metallic materials where there is a large density of states at the Fermi level. This could contribute largely to the heat capacity and ultimately the Gibbs free energy. With these warnings in mind, as well as the use of benchmarks to experimental data to ensure agreement for similar systems, the Debye-Grüneisen model presents a highly valuable contribution to computational studies where vibrational properties are commonly ignored.

Chapter 3

Uncertainty Quantification in Density Functional Theory

Much of this thesis relies on predictions either directly from Density Functional Theory or from models that are trained on Density Functional Theory. To understand these predictions and to understand the relative uncertainty of a model trained on DFT (which can be quantified rather easily) with the uncertainty of the underlying DFT itself (which is not straightforward), we employ an empirical method of uncertainty estimation. We then propagate the uncertainty estimate from the DFT energy calculations to derived properties including thermodynamics predictions.

As mentioned previously, Density Functional Theory has been largely successful in first-principles predictions of materials and their properties and as a result, it is at the core of computational materials design and discovery. Although there are many variations in the practical implementations to improve numerical stability and convergence through projection onto wavefunction basis sets or to save computational time through the approximate treatment of core electrons through pseudo-potentials, it has been demonstrated that the exchange-correlation functional is the leading cause of error across many implementations for well-converged DFT calculations with sufficient parameters related to the numerical precision of the calculation.[43] The Generalized Gradient Approximation described earlier, represents a balance of computational efficiency and accuracy, but given the mathematical freedom of implementing the gradient dependency, a range of GGA-level functionals have emerged, with some more successful for certain use cases than for others [44]. Additionally, two schools of thought exist for the selection and development of exchange-correlation functionals. The first is the reductionist approach in which the functional is crafted to obey known asymptotic and limiting conditions. The contrast to this is a purely empirical approach where some or all of these physics-based limiting conditions are relaxed and a functional is trained on experimental or more accurate theoretical data. Therefore, for every prediction made with DFT, the question emerges: What would the prediction be if a different functional were used? Naively, this could be answered by

repeating the calculation for a selection of GGA-level functionals. This approach, however, would be computationally expensive and as there is no systematic way to pick the collection of functionals to be tested, this method would not guarantee accurate results, especially in the case that a GGA-level treatment of the problem is not sufficient.

Previous studies have attempted to quantify the uncertainty of DFT through statistical and/or regression analyses of DFT predictions over large sets of data.[45, 46, 47] Through these studies, the error in density functional theory has been shown to be a combination of both systematic bias and random error.[45, 46, 47] If we take the empirical approach to exchange-correlation energy selection and view the functional merely as a mathematical model approximating reality, we can understand the selection in the context of the bias-variance trade-off. This is the tension between contributions to the expected model error due to systematic error (the bias) and random error (the variance). Within the field of machine learning, ensemble methods have been successfully used to reduce the variance at the cost of slightly increasing the bias of each individual model in the ensemble with respect to the best alternative. In relation to DFT, however, this bias can be quantified from comparison to a data set assumed to be ground truth and relayed as error bars on the prediction. We will discuss a particular class of ensemble methods that are derived from Bayesian Statistics known as Bayesian Error Estimation Functionals.

3.1 Bayesian Error Estimation Functional

When fitting a model M to some underlying data D , we can say there is a probability distribution that a prediction of that model will give a value y at some input point x given by $P(y|x, D, M)$. If we use a Bayesian ensemble of models, we can average over all of the ensembles:

$$P(y|x, D) = \sum_M P(y|x, M, D)P(M|D) \quad (3.1)$$

In order to evaluate this expression, we must understand the probability that a model is a true representation of reality given the data, $P(M|D)$. Conventionally, when a model is fit, the parameters of the model, θ are adjusted so as to minimize the least-squares cost function between the predicted values $\tilde{y}(\theta)$ and the target values y .

$$C(\theta) = \frac{1}{2} \sum_i (\tilde{y}(\theta)_i - y_i)^2 \quad (3.2)$$

This represents the belief that the lower the cost square error is, the more likely the model represents reality and the lowest value then maximizes the likelihood of estimation. This can be more formally derived if the errors of the between the model and reality are assumed to have a Gaussian distribution and the maximum to this

Gaussian likelihood for the prediction can be easily shown to be equivalent to minimizing the sum of squared error. Depending on the data set used to train the model and the shape of the squared error curve within the parameter space of θ , there could exist models with nominally higher squared error for the data on which the model is trained, but that have reasonable and likely differing predictions on unseen data. It is then useful to sample near the minimum of the squared error well in order to systematically probe the set of reasonable models. In this way, the sensitivity of the predictions made by the fit model can be tested with respect to parameter variations. A useful approximation of the model probability in Eq. 3.1 inspired by statistical mechanics is given then by

$$P(M(\theta)|D) \sim e^{-\frac{C(\theta)}{T}} \quad (3.3)$$

where $C(\theta)$ is the cost function of the model given by least squares and T is some tunable temperature to control how much of the model phase space is sampled. The method was first applied to the fitting of interatomic potentials [48] and latter extended to models for the exchange-correlation potential in DFT at the GGA level [49, 50]. Within these works, an empirical model was fit to the underlying data through the conventional minimization of the cost function. The effective temperature to sample parameter space was then tuned such that the standard deviation of the spread of the predictions matched the error of the model with respect to the underlying data. In this way, the accuracy of the fit model is linked to the variance in the ensemble. When implemented for DFT, the fit model for exchange-correlation potential would provide a fully self-consistent calculation of the electron density and corresponding total energy. The ensemble of predictions can then be computed efficiently by evaluating the energy non-self-consistently using the converged electron density from the best-fit model. Extensions of this work have applied this method to higher-order methods of DFT such as the addition of van der Waals forces [51], and meta-GGA where higher gradients of the density are used [52]. Others alternatives for error estimation have tried to address the question of the likelihood of statistical bias in the estimation of error [53] through system-focused reparametrization [54], machine learning [55], and the creation of more specialized functionals [56].

We focus our attention on the implementation of these so called Bayesian Error Estimation Functionals (BEEF) that includes the addition of van der Waals non local corrections (BEEF-vdW) [51]. In this case, the gradient dependence is imposed through the introduction of an enhancement factor, $F[n, \nabla n]$ that is a function of the gradient of the density and a change of variables: $s = \frac{|\nabla n|}{2k_F n}$ where $k_F = (3\pi^2 n)^{-1/3}$ is the local Fermi wave vector. The exchange energy is now approximated as

$$E_x^{GGA}(n, \nabla n) = \int \epsilon_x(n) F_x[s(n, \nabla n)] d\mathbf{r} \quad (3.4)$$

$$\epsilon_x^{GGA}(n, \nabla n) = \epsilon_x(n) F_x[s(n, \nabla n)] \quad (3.5)$$

In the case of BEEF-vdW, another change of variables is made to restrict to the domain from -1 to 1

$$t(s) = \frac{2s^2}{4 + s^2}, \quad -1 \leq t \leq 1. \quad (3.6)$$

The enhancement factor is then written in terms of Legendre polynomials $B_m[t(s)]$

$$F_x^{GGA} = \sum_m a_m B_m[t(s)] \quad (3.7)$$

so that the exchange energy is then given by

$$E^{GGA-x}[n, \nabla n] = \sum_m \int \epsilon_x^{UEG} B_m[t(s)] d\mathbf{r} \quad (3.8)$$

$$= \sum_m a_m E_m^{GGA-x}[n, \nabla n] \quad (3.9)$$

The correlation functional used is a linear combination of the GGA Perdew-Burke-Ernzerhof (PBE) functional [14] and Perdew-Wang LDA [12]. Non local correlation is accounted for using the vdW-DF2 [18] kernel. Thus the full exchange-correlation functional used is give by [51]

$$E_{xc} = \sum_{m=0}^{M_x-1} a_m E_m^{GGA-x} + \alpha_c E^{LDA-c} + (1 - \alpha_c) E^{PBE-c} + E^{nl-c} \quad (3.10)$$

Just as in the previous work, these coefficients are trained to reproduce the training data. This includes molecular formation energies, molecular reaction energies, non-covalent interaction energies, solid-state properties such as cohesive energies and lattice constants, and energies for chemisorption on solid surfaces all collected from experiment and highly accurate levels of theory. The coefficients α and a_i in Eq. 3.8 can then be varied as given by Eq. 3.3 and evaluated non-self-consistently with the spread tuned to match the error of the best fit functional with respect to the training data. This assumption of non-self-consistent evaluation is well supported by the similarity of the converged BEEF-vdW with other converged densities at the GGA level as seen in by plotting the density of states for a sample material in Fig. 3.1. Furthermore, the error estimation of BEEF-vdW has been shown to bound other general gradient approximation functionals for the prediction of mechanical properties [57], magnetic ground states[58], vibrational properties[59, 60, 61], and reaction enthalpies for hydrocarbons [62, 63] demonstrating it's ability to intelligently sample physically reasonable parameters.

With this ensemble of functionals, not only do we have a quantitative metric for uncertainty through the correspondence of the standard deviation to the error bars on the prediction, but we now have a systematic way to sample GGA space

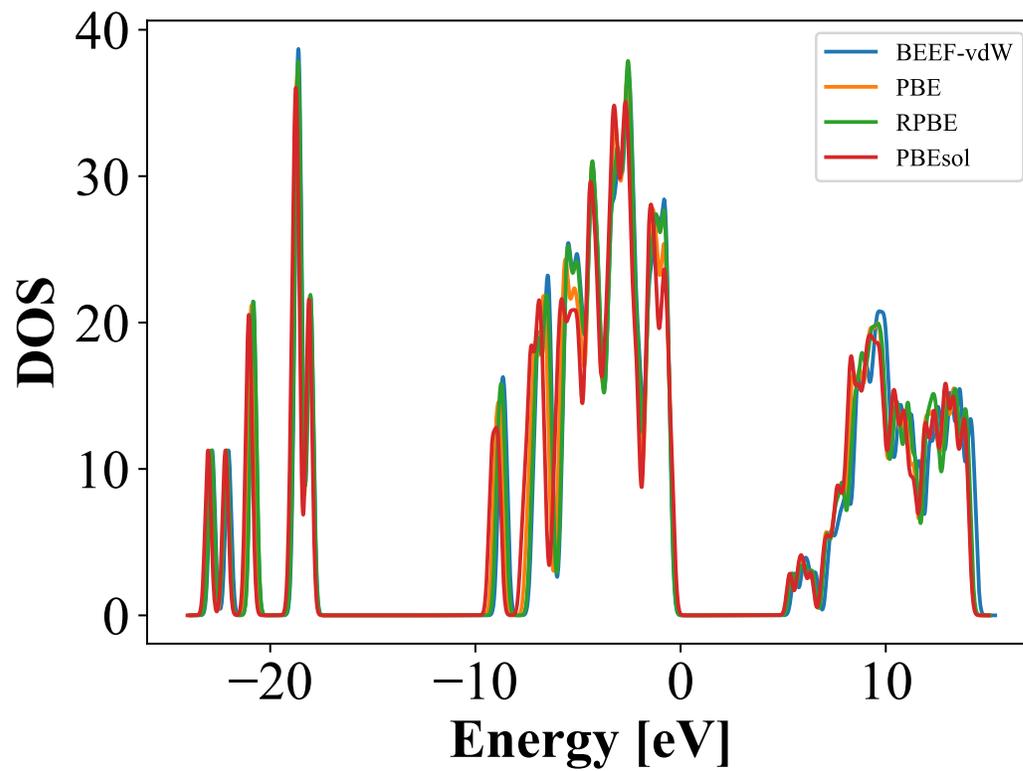


Figure 3.1: The total density of states ferromagnetically aligned LiFePO₄F as calculated by BEEF-vdW, PBE, RPBE, and PBEsol. [Reprinted with permission from Ref. [58]. Copyright 2017 American Physics Society]

and understand the sensitivity of predictions on the choice of exchange-correlation functional. We will now talk about one way to use this ensemble to compute a quantitative metric of prediction confidence in DFT.

3.2 Prediction Confidence

The purpose of this section is to introduce a prediction confidence metric and demonstrate its usefulness through the application towards a problem intrinsic to physics directly related to the exchange-correlation energy: the magnetic properties of materials arising from the exchange energy of between electrons of magnetic ions.

Given the self-consistent nature of DFT, a collection of metastable states can be reached [64] and often these various metastable states are search to understand which of the states is the true ground state. We can predict this ground state by calculating the energy of each configuration of the system, Ω , using DFT and finding the lowest energy

$$E_{GS} = \min_{\Omega}[E(\Omega)] \quad (3.11)$$

Although the energies of these various states may be very close, the states may be qualitative very different, such as adsorption site in surface catalysis, the relative orientation of spins in magnetic materials, or crystal structure and orderings atoms within a material for a given composition. When using a single functional, we again can wonder about the change in the prediction with the functional choice. Here we use the ensemble of energetic predictions of BEEF-vdW to define a prediction confidence as the normalized number of functionals that agree with the predicted of the ground state configuration, Ω_{GS}

$$c = \frac{1}{N_{ens}} \sum_i^{N_{ens}} \prod_j \Theta(E_{\Omega_j,i} - E_{\Omega_{GS},i}) \quad (3.12)$$

In this equation we iterate over all functionals, i , and all other configurations, j , N_{ens} is the number of ensemble functionals and $\Theta(x)$ is the Heaviside step function. For most cases, the distribution of energies are well described by Gaussian distributions and therefore this confidence value can be expressed as

$$c \approx \prod_j \int_0^{\infty} \frac{dx}{\sqrt{2\pi\sigma_j^2}} e^{-\frac{(x-\mu_j)^2}{2\sigma_j^2}} \quad (3.13)$$

It should be noted, that although this confidence value is based on a probability-like treatment of the errors, it makes no prediction of physical probabilities like that of statistical mechanics. Instead, it is a measure of agreement in computations at a

given level of approximation and therefore is loosely related to the probability that a prediction is correct.

We will now look more closely at the prediction of magnetic ground states, which provides a simple yet useful test case for using this prediction confidence given the origin of magnetism in materials. In Chapter 2, we discussed briefly how the formalism of DFT can be extended to collinear spin-polarized systems. In this treatment, the total magnetization of the system can be computed by integrating the difference in spin up and spin down densities to giving the z-projection of the spin-only magnetic moment. Additionally, from this simple spin-density functional theory formalism, a Stoner model for ferromagnetism can be derived[65, 66, 67, 68]

$$E^+ = \epsilon - I \frac{n^+ - n^-}{n}, \quad E^- = \epsilon + I \frac{n^+ - n^-}{n} \quad (3.14)$$

where ϵ is the energy of the unpolarized electrons and I is the exchange energy, the value of which determines whether or not there is spontaneous polarization in the system. Thus, the accuracy of the description of the exchange energy, which is sampled by BEEF-vdW, will determine the magnetic properties of the material.

We benchmark the accuracy of BEEF-vdW in Table 3.1 and find good agreement with respect to a collection of other functionals and experiment. The notable exception, however, is a systematic overestimation of the magnitude of the magnetic moment for Cr, which is well known to have an incommensurate spin-density wave ground state [74, 70]. The agreement on the magnitude of the magnetic moments between functionals is likely due to the similarity of the converged density over all the functionals. The density of states predicted from DFT with various functionals shown in Fig. 3.1 demonstrates this well and supports the assumption of evaluating the energies of the BEEF-vdW ensemble non-self consistently, as we will do later.

To utilize the prediction confidence metric, we explored the magnetic ordering of a collection of materials, testing the ferromagnetic (FM) and antiferromagnetic (AFM) configurations for each material using DFT with the BEEF-vdW functional as well as a range of other GGA functionals in order to probe the ground state spin arrangements. This collection of materials spans a range of crystal structures, complexity, elements, and origin of magnetism. In the case of LaMnO_3 and CuCr_2O_4 , lattice distortions play a large role [75, 76]. The elemental state of Ni and Fe demonstrate a more basic direct exchange ferromagnetic nature [77], while antiferromagnetism (AFM) due to superexchange is seen in the orderings in the oxides tested: FePO_4 , LaMnO_3 , and CuCr_2O_4 . For each material tested, the lattice constants and internal atomic positions were optimized by relaxing to a max force of 0.01 eV/Å. A graphical representation of the confidence value can be seen for Cr in Fig. 3.2. In this figure, the ensemble of energy differences between a nonmagnetic low spin state and the antiferromagnetic state are shown. The shaded region shows the area under the probability distribution curve that sums up all functionals that agree in the prediction that the AFM is more stable.

Table 3.1: Table reproduced from Ref. [58] showing both the predicted and experimentally measured magnetic moment per magnetic ion, μ_B . Both the olivine (o) and quartz (q) polymorphs are included for LiFePO_4 . In the case of CuCr_2O_4 , with two magnetic species, the total magnetic moment is given. Experimental references are from (a) Ref.[69], (b) Ref. [70], (c) Ref. [71], (d) Ref. [72], and (e) Ref. [73]

	Fe	Cr	Ni	$\text{FePO}_4\text{-q}$ $\text{FePO}_4\text{-o}$	LaMnO_3	CuCr_2O_4
BEEF	2.33	1.62	0.61	4.29 4.03	3.89	5.00
PBE	2.13	1.23	0.60	4.31 4.00	3.85	5.00
RPBE	2.21	1.77	0.61	4.33 4.02	3.89	5.00
PBEsol	2.01	0.74	0.58	4.30 3.97	3.77	5.00
Expt.	2.22 ^a	0.62 ^b	0.61 ^a	4.53 ^c 4.02 ^d	3.70 ^e	0.39 ^a

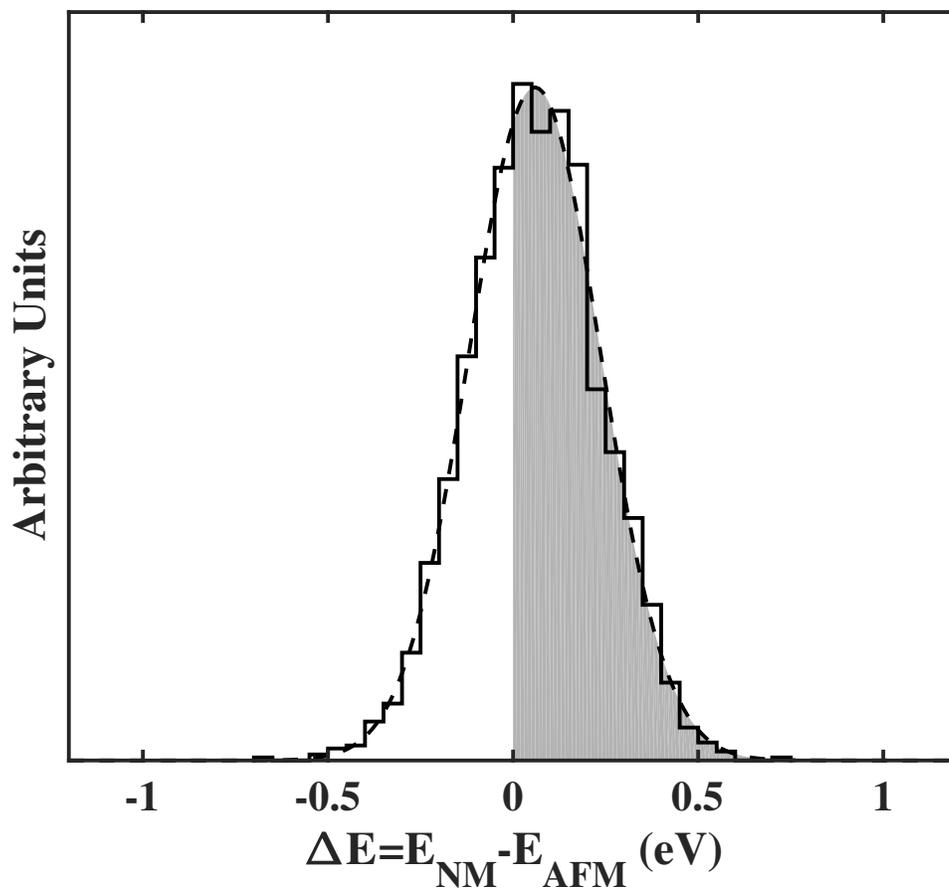


Figure 3.2: The ensemble of energy difference predictions to determine the magnetic ground state of Cr is shown. The c -value for the prediction of AFM as the ground state is shown as the shaded region. Additionally, the normal distribution with ensemble mean and standard deviation is also shown and in good agreement with raw data. [Reprinted with permission from Ref. [58]. Copyright 2017 American Physics Society]

We see in Table 3.2, for most materials tested, both good agreement with experiment and high prediction confidence. This high prediction confidence is echoed by the agreement over the other GGA functionals tested in the sign of $\Delta E = E_{FM} - E_{AFM}$. In the case of Cr, however, we see low prediction confidence. This is likely due to the ground state, which was mentioned previously to be a spin density wave and thus is pointing to a deficiency of the treatment of Cr at the level of GGA. Interestingly, extensions of DFT beyond the semi-local approximation have been suggested as necessary to properly describe Cr or any spin density wave material [78].

Another interesting problem is the ambiguity of AFM orderings. Many materials exhibit multiple length scales over which magnetic alignment or anti-alignment can occur leading to multiple possible AFM states. Many studies, however only investigate the energy difference between FM and a single possible AFM state. While this can determine whether or not the material is FM, in the case that the material is predicted to not be FM, it does not determine the kind of AFM configuration present. We look at two materials with applications as next-generation cathode materials with high capacity. The materials LiFePO_4F and NaFePO_4F have been predicted to cycle up to two alkali ions $\text{Li}_2\text{FePO}_4\text{F}$ and $\text{Na}_2\text{FePO}_4\text{F}$. In a previous first-principles study of these materials and the corresponding magnetic ground state, Ramzan *et al.* found disagreement between the GGA level PBE functional and PBE with an on-site Hubbard U (PBE+U) correction of $U=0.95\text{eV}$ and $J=0.95\text{eV}$ [83]. Additionally, there was no description for the exact AFM ordering being tested. These materials which have an tetragonal conventional cell, have three distinct interaction lengths over which the Fe atoms could magnetically orient. To further explore this disagreement, we repeated the magnetic ground state procedure described above for testing the confidence in the prediction of the magnetic ground over all other possible magnetic configurations. We tested both FM and AFM coupling over the three smallest interaction lengths of the Fe atoms leading to eight total spin states tested for each. We also are able to create an isolated c-value testing the confidence that the coupling is AFM in each particular length scale. The results for these predictions are seen in Tab. 3.3. In the case of LiFePO_4 and NaFePO_4F , where there was previously seen agreement between PBE and PB+U, we see also predict a high confidence value when considering all states versus the one predicted by the main BEEF-vdW functional to be the ground state. For $\text{Li}_2\text{FePO}_4\text{F}$ and $\text{Na}_2\text{FePO}_4\text{F}$, however, we recreate the uncertainty of the predictions with a c-value of 0.614 and 0.6795 respectively. This again points to a possible deficiency in the treatment of these systems at the level of GGA, and suggests the need for semi-empirical Hubbard corrections in these materi-

¹Ref. [79]

²Ref. [70] This is a difference between nonmagnetic (NM) and AFM

³Ref. [80]

⁴Ref. [71]

⁵Ref. [81]

⁶Ref. [73] The AFM state is A-type

⁷Ref. [82] This is a difference between two AFM states.

Table 3.2: Predictions for the magnetic ground state and magnetic energy difference $E_{FM} - E_{AFM}$ of various materials reproduced from Ref [58] the standard deviation, σ , and mean, μ , fro the BEEF-vdW ensemble is given along with the c-value incorporating all magnetic states, as well as the c-value derived from a Gaussian fit of the ensemble.]

Material	BEEF-vdW	PBE	RPBE	PBEsol	exp	c-value	Approx. c-value
Fe BCC $\text{Im}\bar{3}\text{m}$	-0.51 ± 0.20	-0.34	-0.48	-0.41	FM ¹	0.997	0.994
Cr BCC $\text{Im}\bar{3}\text{m}$	0.06 ± 0.18	0.03	0.06	0.01	AFM ³	0.630	0.631
Ni FCC $\text{Fm}\bar{3}\text{m}$	-0.07 ± 0.03	-0.06	-0.06	-0.06	FM ³	0.998	0.997
FePO ₄ α -quartz $\text{P}\bar{3}_1\text{21}$	0.02 ± 0.01	0.02	-0.01	0.00	AFM ⁴	0.952	0.962
FePO ₄ olivine Pnma	0.03 ± 0.011	0.03	0.02	0.05	AFM ⁵	0.999	0.998
LaMnO ₃ perovskite Pbnm	-0.07 ± 0.02	-0.25	-0.11	-0.05	FM ⁶	0.994	0.998
CuCr ₂ O ₄ spinel $\text{I4}_1/\text{amd}$	-0.10 ± 0.06	-0.07	-0.11	-0.05	AFM1 ⁷	0.965	0.969

Table 3.3: The c-values for testing the hypothesis that the magnetic ordering is AFM over the selected length scale in the material. The last row for each material is the c-value testing the prediction of the most stable AFM ordering compared to all eight magnetic orderings tested. Also shown are the results from [83] for PBE and PBE+U with an onsite Hubbard correction of $U=0.95\text{eV}$ and $J=0.95\text{eV}$. The lengths given are the distances between Fe atoms in the FM structure and approximate the lengths between Fe atoms in the different AFM structures.

Material	Interaction Length (\AA)	c-value	PBE	PBE+U
LiFePO ₄ F	3.0	0.777		
	5.8	0.996		
	6.4	0.555		
	all	0.914	AFM	AFM
Li ₂ FePO ₄ F	3.0	0.00		
	5.8	0.216		
	6.6	0.785		
	all	0.6145	FM	AFM
NaFePO ₄ F	3.6	1.00		
	5.9	1.00		
	6.7	0.500		
	all	0.812	AFM	AFM
Na ₂ FePO ₄ F	3.5	0.122		
	6.0	0.733		
	6.9	0.096		
	all	0.6795	FM	AFM

als. This is unsurprising as the well studied un-fluorinated analog of these materials LiFePO₄, has been shown to require a U correction to properly capture the charge localization that leads to the phase separation seen experimentally [84].

This concept of confidence value has been used to understand the prediction confidence for predictions in catalysis [85], surface phase diagrams [86], design principles for 2D functional materials [87], and Li ordering in battery electrodes [88, 89].

We close with a quick discussion of the limitation of this method of uncertainty quantification. A well-known puzzle in catalysis is the systematic failure of DFT at all levels to predict the correct adsorption site of CO intermediate on PT 111 surfaces [90]. DFT predictions systematically predict the incorrect favorability of the adsorbate to prefer the hollow site over the top site. Even with the addition of vdW corrections or when moving to a higher level of theory with mGGA, the problem

remains [91]. A quick analysis with the BEEF-vdW functional provides a c -value for the incorrect prediction of 0.86, which while not overwhelmingly high, supports the widespread inability of DFT to correctly describe the system. Recently it was shown that by using the density converged with an on-site Hubbard-U correction for self-interactions with the PBE functional (PBE+U), evaluated non-self-consistently with the PBE functional, the correct site could be predicted [92]. This suggests that the error in the predicted electron density and not in the exchange-correlation function is the core of this puzzle. Unfortunately, the ensemble error estimation presented here as well as many other schemes for uncertainty quantification currently cannot quantify the uncertainty in the electron density itself.

In summary, we show a rigorous way of applying the ensemble of energy predictions generated by the BEEF-vdW functional to understand the level of certainty in the qualitative prediction of ground state configuration for materials. While many cases gave high prediction confidence and good agreement with experiment, which highlight the success of DFT at the GGA level, we saw notable cases where the disagreement illuminated by this confidence value suggested a fundamental limitation to the treatment of the problem. This method of uncertainty quantification takes the understanding of error bars to a level where we can understand not only a numerical uncertainty but answer the question: If another functional was chosen at the GGA level for this problem, would the science that is predicted change?

3.3 Uncertainty Quantification in Thermodynamic Predictions

We now move to another application of uncertainty quantification from DFT. Here we attempt to understand how the error propagates through physical models that use DFT as input. The energies from DFT represent zero temperature predictions. For many materials, however, finite temperature properties play a large role. In particular, the vibrational properties of crystals can contribute largely to the entropy and free energy and therefore change the predicted stability of structural phases. The Debye-Grüneisen model detailed in Chapter 2 can provide relatively accurate predictions of these vibrational properties by taking the simple approximations of the original Debye model and adding corrections in the form of a scaling factor for sound velocities and the Grüneisen parameter for anharmonic effects. We use this model as a test of the sensitivity of exchange-correlation functional on the predicted finite temperature properties of crystals.

Conventional methods of understanding the propagation of uncertainty through a physics based model assume simple distributions for the uncertainty that are then propagated forward [93]. This assumes, however, that the error from a functional in the prediction of one energy of a material is independent of the error in the prediction of the energy of the same material at for example a different volume. It is well

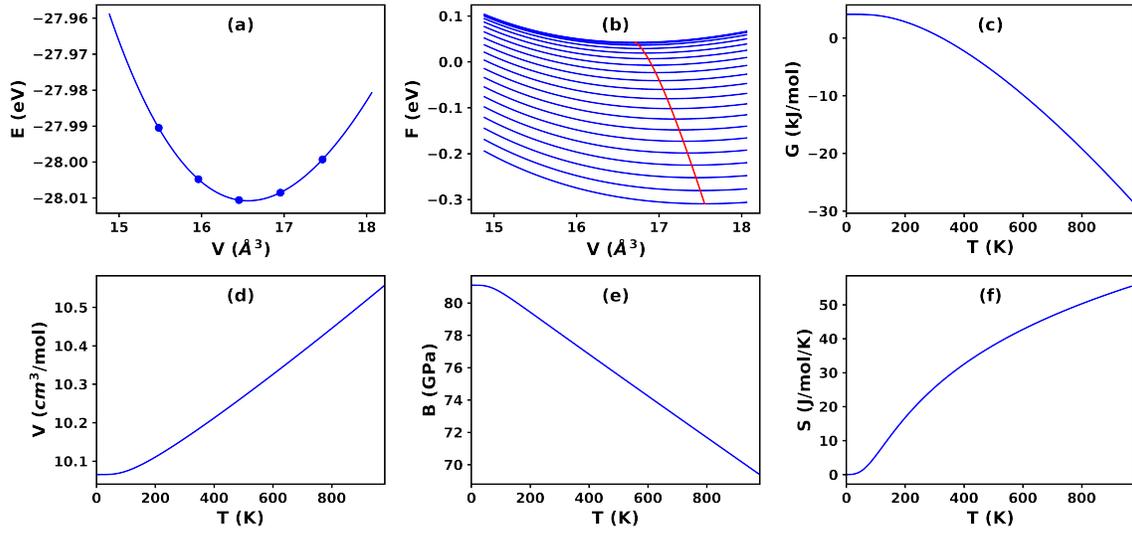


Figure 3.3: Schematic representing the procedure to calculate thermodynamic properties from the Debye model: First, an E-V curve (a) is fit and the free energies at different temperatures are calculated (b). The red line consists of the minimum of free energy at each temperature, from which the Gibbs energy (c), the volume (d), the bulk modulus (e), and vibrational entropy (f) are then computed. [Reproduced from Ref [59], with the permission of AIP Publishing, Copyright 2019]

understood, however, that the errors for the prediction of very similar materials or different compositions in various configurations with DFT are correlated. As the prediction of thermodynamic properties ultimately relates to not only the energy at a given volume, but how the energy changes with respect to volume, and therefore energy differences, this correlated error then becomes an important factor in the propagation of error in thermodynamic modeling. In our approach, by utilizing the ensemble of energies from the BEEF-vdW functional as the statistical data generation step, we can continue to correlate the predictions from each functional at every stage in the process. This thus allows for proper treatment of the statistical dependence of the predictions and the correlated error of DFT. As this correlated error usually results in error cancellation, neglecting the statistical dependence will result in an over prediction of uncertainty.

3.3.1 Application of the Debye-Grüneisen Model

The energy of the material at various volumes around the minimum volume is predicted using DFT. From these energies, a Vinet equation of state is fit [94]. From this, the equilibrium volume, V_0 , the bulk modulus, B_0 , and the derivative of the bulk modulus with respect to pressure, B' can be extracted. As discussed in the

Methods Section from Equation 2.55 for the Debye temperature and Equation 2.42 for the vibrational contribution to the free energy, all vibrational properties can be computed under the Debye-Grüneisen approximation using only these quantities. The Helmholtz free energy binding curves are then computed at various temperatures as seen in Fig. 3.3 (b). From each of these curves, the minimum volume is extracted. When calculated the Gibbs energy,

$$G(P = 0, T) = \min_V F(V, T) + PV \quad (3.15)$$

the pressure effect is small enough to ignore. We, therefore, approximate the Gibbs energy as that of zero pressure

$$G(P = 0, T) = \min_V F(V, T) \quad (3.16)$$

From here, all thermodynamic properties of interest can be derived, including the entropy, $S = -\frac{dG}{dT}$, the enthalpy, $H = G + TS$, the isobaric specific heat, $C_p = \frac{dH}{dT}$, and the volumetric thermal expansion coefficient, $\alpha = \frac{1}{V} \frac{dV_0}{dT}$. In these cases, finite difference is used for all derivatives of with respect to T.

This process was performed for a collection of materials including Mg, Ca, Al, Li, GaAs, NiO, Li₂O, and Al₂O₃ [59]. For each material, each property was predicted using a collection of GGA level functionals self consistently, as well as the full collection of BEEF-vdW functionals non-self-consistently. We highlight the results for Li in Fig. 3.4 as it not only provides an interesting case for understanding the sensitivity of errors but is relevant to the prediction of battery properties later. For Li, as well as all other materials tested, we see rather good agreement in the prediction of Gibbs energy, with variations on the order of 0.01 eV/atom, although the variation increased with temperature. We also see very good agreement between functionals as well as with experiments in the prediction of heat capacity, C_P . When predicting higher-order quantities, such as Bulk modulus, and thermal expansion coefficient, we see significantly more error. For the thermal expansion coefficient, however, we see that although the majority of the data is very close to the mean, some outliers exist extremely far away. This is highlighted by plotting the log of the BEEF-vdW probability density for the thermal expansion coefficient in 3.4 so the few outliers can be seen clearly. This suggests that for most reasonable exchange-correlation functionals, a similar α is predicted, while a functional that is just slightly worse than average will have a huge error. We also see a systematic overestimation of the thermal expansion of Li for all GGA functionals tested. This could be due to both the anisotropy in the bulk properties of lithium and the inelastic behavior of lithium due to its softness.

For the purpose of this Thesis, we also investigate the thermodynamic predictions and associated uncertainties of LiCoO₂ in Fig. 3.5. We immediately see better agreement than Li amongst all functionals and with experiments for the thermal expansion coefficient. But as in the case of Li, we also see few outliers very far from

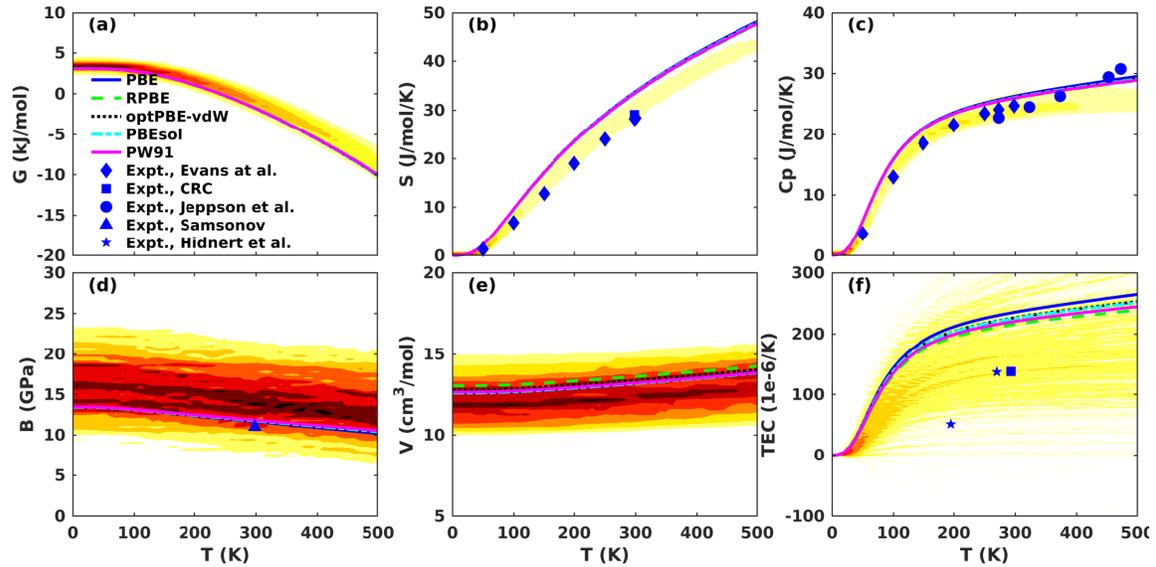


Figure 3.4: Prediction of thermodynamic properties (a) Gibbs energy (b) entropy (c) isobaric heat capacity (d) bulk modulus (e) volume and (f) volumetric thermal expansion coefficient for Li using the Debye-Grüneisen Model compared with experimental measurements for Ref.[95, 96, 97, 98, 99]. For each property, the probability distribution from the BEEF-vdW ensemble is given. In (f), the natural log of the probability density is shown to demonstrate the extreme outliers. [Reproduced from Ref [59], with the permission of AIP Publishing, Copyright 2019.]

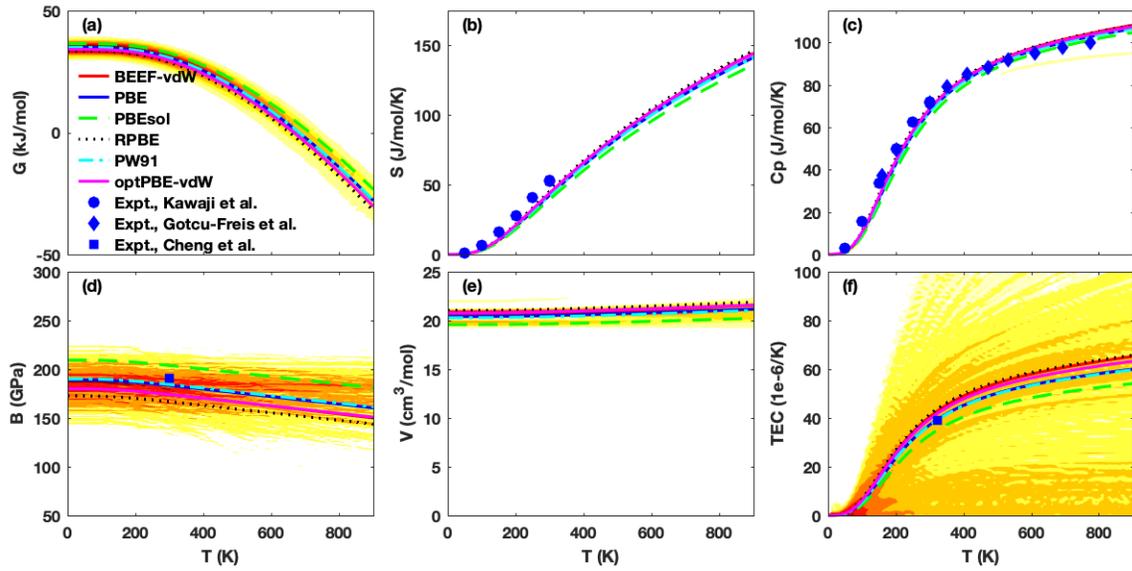


Figure 3.5: Prediction of thermodynamic properties (a) Gibbs energy (b) entropy (c) isobaric heat capacity (d) bulk modulus (e) volume and (f) volumetric thermal expansion coefficient for LiCoO_2 using the Debye-Grüneisen Model compared with experimental measurements for Ref. [100, 101, 102] cite, cite, cite. For each property, the probability distribution from the BEEF-vdW ensemble is given. In (f), the natural log of the probability density is shown to demonstrate the extreme outliers.

the typical prediction, supporting the high level of sensitivity on the input energies. Just as with Li, the entropy and heat capacity are well captured with comparison to experiment for all functionals tested.

3.4 Summary

In this chapter, we explored the uncertainty from the choice of the exchange-correlation functional and how it can affect DFT predictions. We first discussed a semi-empirical way to understand DFT uncertainty through the training of an exchange-correlation functional with respect to experimental and high accuracy quantum chemistry data. Using Bayesian statistics, this can then generate an ensemble of predictions that are near the best fit functional in the model parameter space. A framework for understanding the prediction confidence for ground-state configurations was introduced and employed using the ensemble of energies from the BEEF-vdW functional. This framework was applied to magnetic ground states due to the origin of magnetic coupling in the exchange energy of electrons. We found good agreement for systems previously thought to be well described by DFT and for select interesting cases found that low prediction confidence may suggest a higher level of theory or semi-empirical corrections are needed. We then applied the BEEF-vdW ensemble to understand the variation in Debye model predictions. This method has the added benefit over other methods of uncertainty quantification in properly treating the correlated errors of a set of predictions from a single functional, which are treated as independent in other analyses. We found good agreement with respect to experiment for finite temperature predictions for most properties of interest in Li and LiCoO₂, two materials of interest to this Thesis. We did, however, see that properties related to higher-order derivatives are more sensitive. This emerges in the prediction of the thermal expansion and therefore the thermal expansion coefficient could be used to understand how accurate the energy-volume prediction is when we train a machine learning potential later.

Chapter 4

Transition Metal Cathodes

This chapter focuses on applying the methods and techniques discussed in Chapters 2 and 3 to battery cathodes. Intercalation cathodes generate usable electrochemical energy through the reaction with some intercalant, usually an alkali metal with a small radius and large electropositivity. As a result, the most desirable intercalant is thus Li. The cathode not only provides a host structure for the lithium-ion to be inserted but also an electronic state for the lithium's electron to go to. We can, therefore, think of this reaction as lowering the chemical potential of the lithium with respect to what it was in the elemental state. As seen in Fig. 4.1, the lithium-ion travels through the electrolyte and the electron through the external circuit. The lithium-ion is then intercalated into the host material, while the electron is absorbed into the redox couple of the transition metal within the cathode. During charge, the lithium is then pulled out of the cathode with an applied potential. In practice this reaction and its reversibility are imperfect. The amount of lithium extracted and replaced and the potential range over which the battery is cycled can be varied and controls not only initial capacity and voltage but more importantly the long term performance. Given the need for electrification of transportation which requires increasingly higher capacity and cycle life, a detailed atomistic understanding is needed. When a particular composition is cycled from 3.5 V - 4.2 V, how does that compare to another composition in the same range? What are the lithium compositions at each voltage? All of these questions require the use of the computational techniques explored in this chapter.

The first intercalation electrode LiTiS_2 was verified in 1976 [103] and shown to have a potential of 2V with respect to the Li/Li^+ redox couple. Unfortunately, electroplating of metallic lithium at the time, and to this day, presents major issues to the operation of the battery over multiple cycles. Any replacement of the metallic lithium with another lithium alloy or intercalation compound leads to undesirably low cell voltage. In the search for higher potential cathodes, the transition metal should be coordinated with anions that are both small and largely electronegative. From this analysis, we are led to oxide materials over sulfides as originally tested. Additionally,

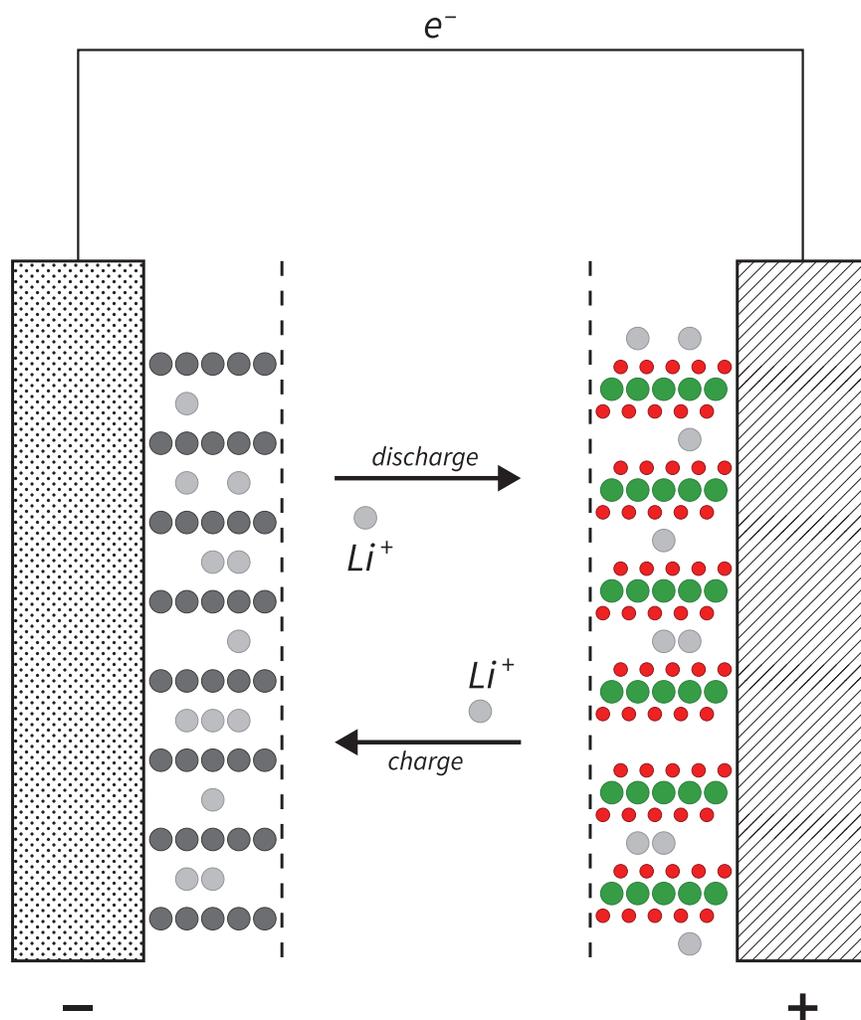


Figure 4.1: Schematic for a lithium-ion battery. On the left, we have the anode (denoted by -) which is conventionally graphite. On the right, we show the cathode (denoted by +) as a layered transition metal oxide. The lithium moves through the electrolyte from the anode to the cathode during discharge as the electron goes through the external circuit to provide usable electricity. During charge, an applied potential moves this process in reverse. The lithium moves out the cathode and back into the anode.

the metal should be in a high oxidation state to further lower the energy of the lowest unoccupied state.

In 1980, the layered structure of $LiCoO_2$ was first shown to reversibly cycle lithium with a cell potential of near 4V [104]. This material, later commercialized in 1991 by Sony, revolutionized the market of rechargeable batteries. This layered structure, seen in Fig. 4.2 enables high lithium mobility allowing for nearly all of the lithium to be cycled. In practice, however, only half of the lithium is removed from the struc-

ture corresponding to a maximum potential of 4.2V and a capacity of only about 135 mAhg⁻¹. The material is shown to have poor capacity retention when more Li is extracted by cycling up to 4.5V. This has been attributed to a monoclinic phase transition at Li_{0.5}CoO₂ that is accompanied by a large expansion in the layer separation [105]. Others have suggested these structural changes are only relevant above 4.5V and that surface reactivity and reconstruction lead to increased impedance and therefore the poor performance seen [106]. The use of surface coatings has been shown to improve the performance of LiCoO₂ [107]. Interestingly, the use of surface coatings would address both possible issues by mechanically suppressing the lattice expansion and contraction as well as stabilizing the surface and preventing the suspected reactions with the electrolyte. As an alternative cathode, LiNiO₂ in the same layered form was studied in the early 1990s and shown to have a much larger usable capacity nearing 200 mAhg⁻¹ and a reversible capacity of about 150 mAhg⁻¹ at the expense of increased synthesis difficulty due to increased disorder of Li/Ni swapping [108]. This larger capacity, however, quickly reduces over cycling due to lattice cracking from a monoclinic distortion near $x_{Li} = 0.5$ and rapid layer contraction at high states of charge [109, 110]. Also of note, is the layered LiMnO₂ which was successfully synthesized but showed rapid capacity loss directly linked to an irreversible transition to spinel LiMn₂O₄ [111].

The success of Co in this layered structure could inspire the use of other transition metals in the same structure possible beyond the ones already mentioned. The formation of this layered structure occurs when the M³⁺ oxidation state of the metal has an ionic radius that is much smaller than that of Li⁺ as the bond lengths of the metal and lithium-oxygen octahedra can then relax independently within the material. This occurs in the case of Co, V, Ni and Cr. When the cation is the same size as Li⁺, such as in the case of Fe or Sc, the material will then form an ordered structure where the metal and lithium share each of the layers equally. The lithium is therefore trapped in the structure and cannot be easily extracted for electrochemical cycling. The much cheaper and seemingly promising layered LiFeO₂ can be synthesized but show poor electrochemical performance as a result of this [112]. In the desired layered structure, as the lithium is removed the lattice will become increasingly unstable which is detrimental to the long-term cycling of these materials as battery electrodes. The cubic closed packed (ccp) structure of the oxygen anion in the layered structure is identical to that seen in the spinel structure. In the spinel structure, cations occupy both the octahedral and tetrahedral sites. Therefore a phase transformation to this spinel phase or to the rock salt phase would remove usable capacity of the cathode. To prevent the migration of the transition metals to the Li vacancies driven by electrostatics, or to tetrahedral sites to form a spinel phase, we look for transition metals with a high preference to stay in the octahedral site as this will correspond to a large kinetic barrier to move through the tetrahedral site that is on the way to the lithium vacancy octahedral site [113] as well as prevent spinel formation. This preference must also be preserved as the metal oxidizes. Previous calculations of the

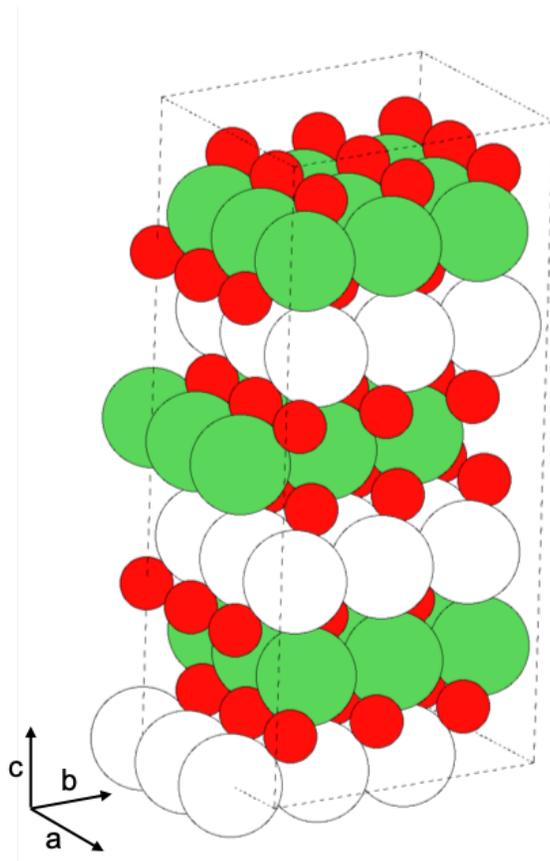


Figure 4.2: The O3 structure of $\text{LiNi}_{1-x-y}\text{Mn}_x\text{Co}_y\text{O}_2$ with rhombohedral symmetry given by the $R\bar{3}m$ space group. The transition metals are shown in green and form a triangular lattice. The oxygen in red form a cubic close-packed like structure. And the white show Li or lithium vacancies. These together also form a triangular lattice. The transition metals and the lithium sites are octahedrally coordinated with the oxygens leading to stacking of the layers such that they repeat every three layers, giving the O3 name.

relative energies of transition metals in the 3+ and 4+ states in the octahedral and tetrahedral site of the ccp oxygen lattice leave only Co and Ni as candidate transition metals [113]. In the case of Mn, the 4+ oxidation state present in highly charged, low lithium content Li_xMnO_2 is very stable. On the other hand, the 3+ state seen in lower states of charge and high lithium content will disproportionate quickly to 2+ and 4+ followed by a migration of the Mn^{2+} to the tetrahedral site to form spinel as was mentioned previously. While pure LiMnO_2 is not viable as a cathode material, the addition of Mn in a layered material containing Ni, however, will create a more stable structure. The Ni^{3+} and Mn^{3+} are known to charge disproportionate to Ni^{2+} and Mn^{4+} , giving stability to the structure given the high barrier of Mn^{4+} to migrate.

It is, therefore, no surprise that the ternary compound $\text{LiNi}_{1-x-y}\text{Mn}_x\text{Co}_y\text{O}_2$, was then studied in 1999 and shown to have improved reversible capacity over LiCoO_2 and improved capacity retention over LiNiO_2 [114]. This material, commonly abbreviated NMCXYZ, where X, Y, and Z are the proportions of Ni, Mn, and Co respectively (e.g. $\text{LiNi}_{0.8}\text{Mn}_{0.1}\text{Co}_{0.1}\text{O}_2$ is written NMC811), has been remarkably successful in providing high capacity, high energy density batteries. The most common proportion NMC111 was first synthesized in 2001 [115]. This material emerged as a leading candidate cathode material as it was the first cathode material to deliver a long term reversible capacity closer to 200 mAhg^{-1} and showed remarkably more stability against reaction with the electrolyte at Li contents below 20% than LiNiO_2 and LiCoO_2 [116]. Today, NMC occupies the majority of the EV market share yet may still provide the possibility for actionable improvement [1]. The desire to increase in capacity as well as the socio-economic and supply concerns surrounding Co [2], has to lead to increased interest in high-Ni containing NMC. This has been reflected in the development towards the commercialization of high-Ni NMC811. As the nickel content increases, however, as with pure LiNiO_2 , we see lower capacity retention [3] and increased risk of dangerous degradation due to increased O_2 evolution and lower temperature stability [117].

Unfortunately, understanding of trends in degradation and performance with respect to compositions is limited to small experimental searches of the full phase space. Additionally, suggestions for improvement are largely limited to single compositions and the solutions for one composition may not apply to the issues of another composition. What is therefore required is a computationally efficient method to understand these materials. Before focusing in more detail on the specifics of the limitations of increased Ni, we turn to understanding the NMC phase space as a whole. It is largely believed that the material space is a solid solution and arbitrary phases can then be synthesized. Experimental studies, however, show negative mixing enthalpy [118], and studies of transition metal ordering of NMC111 have shown at least short-range ordering [119], both suggesting that NMC is not a disordered solid solution at room temperature. This is further supported by the very few compositionally phases used experimentally as seen in Fig. 4.3. Therefore to understand and predict something as fundamental as the lithium ordering and therefore the open-circuit voltage of the

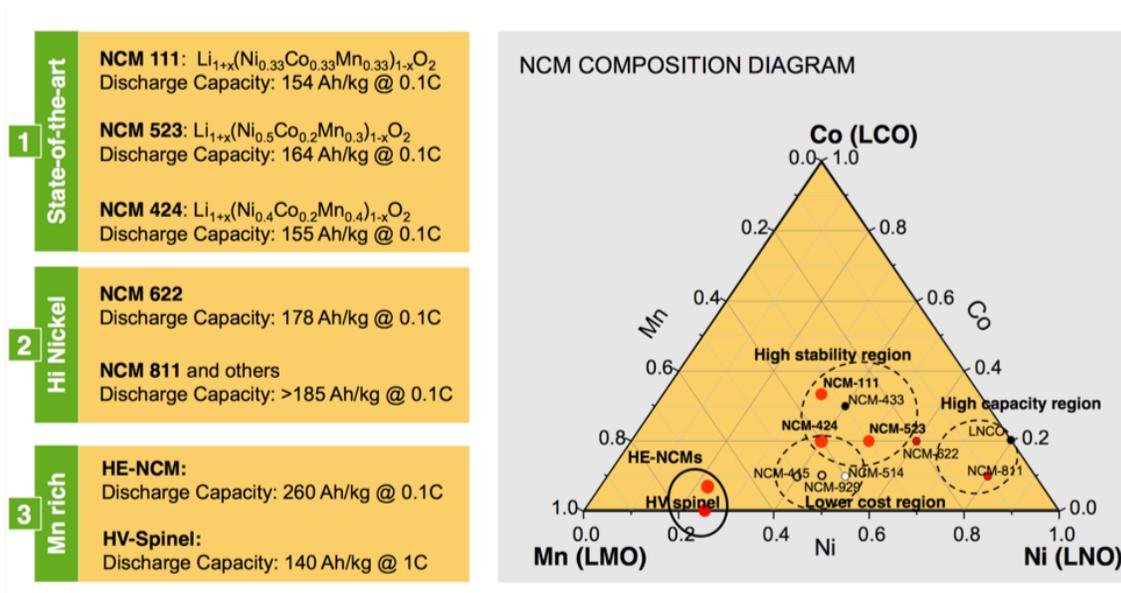


Figure 4.3: A catalogue of conventionally used NMC (referred to as NCM here) compositions. Not the sparsity in the number of compositions used. [Reproduced from Argonne National Lab Copyright 2014.]

battery, the orderings of the transition metals must first be known. It is, therefore, possible that an unfavorable ordering (or full disordering) of transition metals that can be entropically stabilized at synthesis temperatures ($\sim 1000\text{K}$), will be unstable during cycling. At very low Li content, the transition metals all have a tendency to migrate to the vacant lithium sites [120] which could be worsened by a less than favorable transition metal ordering. The transition metals with then attempt to migrate to form a more stable transition metal ordering, degrading the material and diminishing long term capacity. We therefore first seek to understand the extent of this ordering and predict the stable ordered compositions to guide further investigation of NMC compounds.

4.1 Cluster Expansion

For the problem at hand of predicting the set of ground-state configurations and compositions, DFT alone is computationally infeasible. For a fixed lattice size of N , there would be 3^N different materials to screen. While this number is reduced when considering equivalent states due to symmetries, an alternative approach is needed. Reduced-order energetic models trained on DFT data such as cluster expansion have previously been used successfully to describe many materials including transition metal oxide cathodes. [84, 121, 122, 123, 124, 125] In this method, the energy of the material is decomposed into series expansion in sets atoms arrangements called clusters. These clusters could include nearest neighbor, next nearest neighbor, three-body interactions, etc. The number of occurrences of each of these clusters, each of which has a learned contribution to the total energy, is counted and the energy can be approximated. Due to symmetries in the crystal, many of these cluster interactions are equivalent aiding in the training of the model. The occurrence of each cluster is accounted for using occupation variables $\sigma_{i,x}$ that take the value of 1 or 0 representing if species x occupies site i in the lattice or not. The cluster interaction α then carries a weight of J_α . In the case of the NMC energy model for, an Ising like collinear spin interaction term can also be included to account for the magnetic moments of the transition metals. In this case, spin variables taking the value of ± 1 represent spin up or down, and a spin interaction term, K_{xy} , represents either ferromagnetic or antiferromagnetic interaction, depending on the sign, between an atom of species x and an atom of species y . The final model is then

$$\Delta E_f = \sum_{i,x}^N h_x \sigma_{i,x} + \sum_{\langle i,j \rangle, x,y} J_2^{xy} \sigma_{i,x} \sigma_{j,y} + \sum_{\langle i,j,k \rangle, x,y,z} J_3^{xyz} \sigma_{i,x} \sigma_{j,y} \sigma_{k,z} + \dots + \sum_{\langle i,j \rangle, x,y} K_{xy} \vec{S}_{i,x} \cdot \vec{S}_{j,y} \quad (4.1)$$

As only the positions of the transition metals are changing, we choose to treat the oxygen and lithium as a constant background and work on an idealized triangle lattice that the transition metals occupy. A visualization of the idealized lattice used

and an enumeration of the unique cluster configurations can be seen in Fig 4.4. We choose to train the energy model not on the absolute energy of an NMC material $\text{LiNi}_{1-x-y}\text{Mn}_x\text{Co}_y\text{O}_2$, but on the energy difference of the NMC phase mixture with respect to the pure component end members.

$$\Delta E_f = E_{\text{LiNi}_{1-x-y}\text{Mn}_x\text{Co}_y\text{O}_2} - xE_{\text{LiNiO}_2} - yE_{\text{LiMnO}_2} - (1-x-y)E_{\text{LiCoO}_2} \quad (4.2)$$

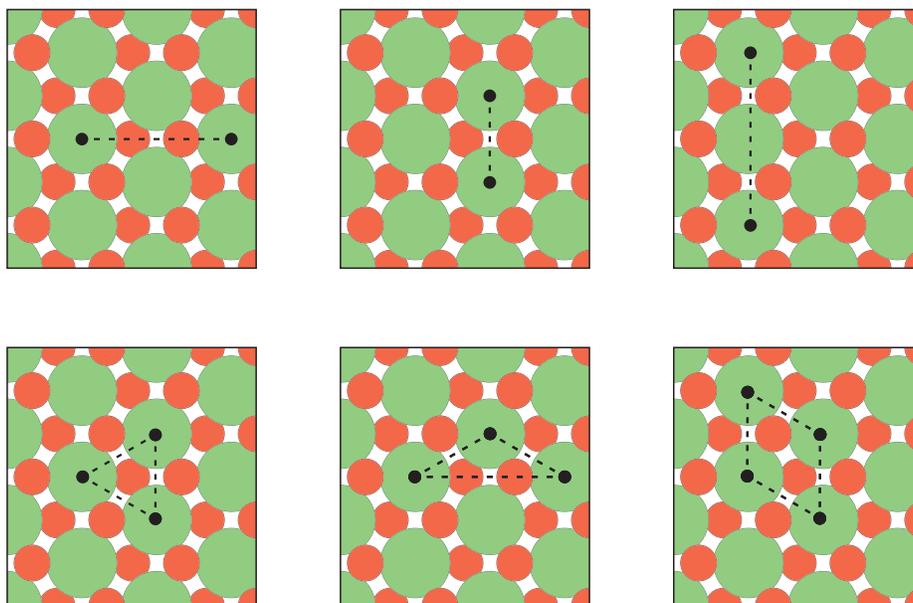
The DFT predictions of energy differences should be more accurate due to the correlated error. Predicting the absolute energy would include the addition of a E_0 mathematical bias term to account for the background energy of the lithium and oxygen lattice that is constant in all of the NMC compositions. The scale of the energy differences ΔE_f and therefore the learned J and K coefficients are on the order of 0.1 eV while the E_0 term would be on the order of 100 eV leading to a model that is hugely numerically unstable. Additionally, since we ultimately care about energy differences to determine favorability of one phase over the other, the errors of the prediction from cluster expansion of absolute energy of two materials to be compared could be compounded to provide a worse final prediction.

It should also be noted that the inclusion of every single cluster interaction would lead to an overdetermined system due to mathematical constraints derived from the formula for composition and lattice symmetries. A set of schemes exist for choosing the complete basis set of cluster interactions to remove linear dependencies. The scheme here is chosen in an attempt to maximize the physical interpretability of the cluster expansion model. From the formula $\text{LiNi}_{1-x-y}\text{Mn}_x\text{Co}_y\text{O}_2$ once the occupancy of two of the species is known, the third species occupation term is fixed, introducing the first linear dependence. Next, the triangular lattice has a fixed number of cluster interactions of each type. For example, there are only 6 total nearest neighbors for each atom, thus the number of nearest-neighbor interactions between Co and another Co can be written in terms of the number of Co on the lattice and the number of Co-Mn and Co-Ni interactions.

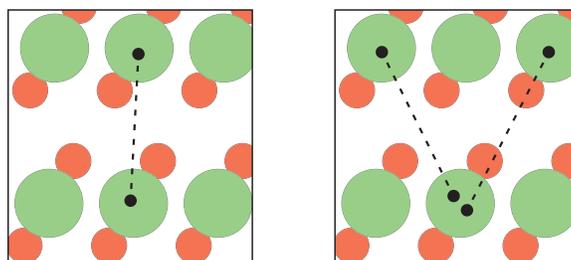
$$N_{Co} = \sum_i^N n_{i,Co}, \quad N_{Co,X} = \sum_{\langle i,j \rangle}^N n_{i,Co,j,X} \quad (4.3)$$

$$N_{Co,Co} = 3N_{Co} - \frac{N_{Co,Mn} + N_{Co,Ni}}{2} \quad (4.4)$$

We, therefore, choose a convention that removes these linear dependencies by rewriting all the homogeneous n-body interactions in the cluster expansion in terms of occupation terms and heterogeneous n-body interactions as demonstrated above. While this may differ from other conventions of cluster expansion, this choice of model provides a more direct understanding of the relative favorability of interaction when mixing species.



(a) In plane cluster interactions from a top view



(b) Out of plane cluster interactions show from side view

Figure 4.4: A cluster interactions considered up to 4 body within a cutoff of 6 Å.

4.1.1 Training Set

To train the cluster expansion model a data set of 118 DFT calculations is used with various compositions, cell shapes that generate the same lattice, and cell sizes up to 120 atoms per cell. The structures were generated using a custom-written code [126] by randomly iterating through composition space and checking the uniqueness of cluster interactions for all clusters shown in Fig. 4.4. While this technique may wrongly classify two structures that are not truly equivalent as the same, it will guarantee each new structure tested is unique compared to all other structures in the data set. The lattice constants of layered NMC vary slightly based on composition so a lattice optimization was performed for each structure. First, the c-direction lattice constant was strained to an $x = \frac{c}{c_0} = 0.9, 0.95, 1, 1.05, \text{ and } 1.1$. The resulting binding curve was then fit to an equation of state [127] and the minimum energy c-lattice constant is found. This is repeated for the a=b lattice constant. The internal atomic coordinates are then relaxed to a max force of $0.03\text{eV}/\text{\AA}$. The final resulting fully relaxed energy is then used to train the model.

4.1.2 Model Selection

One of the main advantages of training and using an energetic model as simplistic as cluster expansion is that only a relatively small DFT training set is needed. This is in contrast to the conventional wisdom of the computer science and machine learning community where large data sets are used to allow for splitting into training and testing data sets to aid in the transferability of the model. We, therefore, need ways to understand the effect of the cluster expansions complexity and number of terms on the balance between the bias and variance of the model. We must make the model as true to the underlying DFT data as possible (low bias) but not overly expressive so as to overfit and not generalize (low variance). A series of model selection techniques applicable in particular to small data set model selection are used here.

The most common validation technique used in cluster expansion is leave one out cross-validation. For a data set of N points, the cluster expansion model is trained on N-1 points and the error is computed on the left-out data point. This is repeated for all of the data points and the root mean squared of all of these errors is reported.

Another way to select a model is to weight the model on both the ability to reproduce the training data and explicitly penalize the model complexity through the number of parameters. This was first done by Hirotugu Akaike in 1974 in what is known by the Akaike Information Criterion[128]

$$AIC = 2k - 2 \ln \hat{L} \tag{4.5}$$

Here \hat{L} is the maximum likelihood estimator for the errors and k is the number of parameters in the model. If we assume the residual errors in the model, $\delta E =$

$E_{DFT} - E_{CU}$ are independent and identically distributed random variables each with a Gaussian distribution, the likelihood function is then given by

$$L = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\sum_i^N (\delta E_i - \mu)^2} \quad (4.6)$$

The maximum likelihood estimator for the mean, μ , and variance, σ^2 , can be found from taking the derivative of this likelihood function with respect to μ and σ and found to be given by

$$\mu = \frac{\sum \delta E_i}{N} \quad (4.7)$$

$$\sigma^2 = \frac{\sum (\delta E_i - \mu)^2}{N} \quad (4.8)$$

Now when the log of this maximum estimator is taken, a simple form emerges due to cancellation and ignoring constant terms.

$$\ln \hat{L} = -\frac{N}{2} \ln(2\pi) - \frac{N}{2} \ln(\sigma^2) - \frac{1}{2\sigma^2} \sum_i^N (\delta E_i - \mu)^2 \sim -\frac{N}{2} \ln(\sigma^2) \quad (4.9)$$

We have ignored the first term as it is only dependent on the number of training points, N , and is constant for all models trained. In the second the sum and σ^2 can cancel out to also give a constant term that is ignored. What is left is the log of a sum of squares like term. For linear regression models, like cluster expansion, the AIC and leave one out cross-validation will asymptotically choose the same model as the number of training points increases [129]. In our case, however, we have far from infinite training points and for small data sets, the AIC has a known tendency to favor overly complex models and a corrected AIC is suggested [130]

$$AIC_c = AIC + \frac{2(k+1)(k+2)}{N-k-2} \quad (4.10)$$

A third and final weighting of model complexity is the Bayesian Information Criterion (BIC). In the formulation of this, it is assumed that within the set of models exists a true model and that the BIC will pick the true model with probability 1 as the number of data points tends to infinity [131].

$$BIC = \ln(N)k - 2 \ln \hat{L} \quad (4.11)$$

Every combination of cluster interactions was tested including all the testing of all 2-body spin interaction. The magnetic moment on the Co atoms converged to a small magnitude and therefore models that both include and do not include Co spin are also test. The results of each of these 4 model selection techniques, seen in Fig 4.5, all chose the same 42 parameter model with next-nearest-neighbor spin, not including

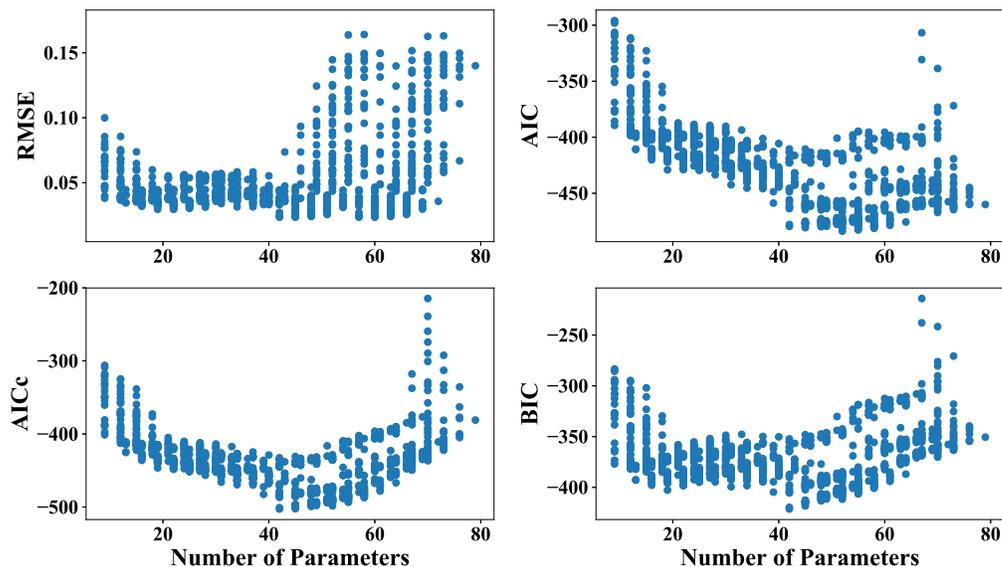


Figure 4.5: Results for the four model selection techniques, (a) leave one out cross validation root mean squared error (RMSE), (b) the Akaike information criterion (AIC), (c) the the corrected Akaike information criterion (AICc), and (d) the Bayesian information critereon (BIC). [Reprinted from Ref [132], Copyright IOP Publishing 2020]

Co spin, 4 body interaction, and out-of-plane next nearest neighbor interaction. This model gave a leave on out root mean squared error (RMSE) of 0.023/eV per formula unit (5.8 meV/atom).

4.1.3 Prediction of Transition Metal Ordering

From our final selected model out of all possible cluster expansion models, we can then predict the ordering of the transition metal cations in the lattice. The layered NMC phase space is largely accepted as a solid solution for most compositions, except high Mn which is unstable and will form spinel. This range is approximately above 40% Mn content at zero Ni content and approximately above 50% Mn content at zero Co content [133]. Both computational and experimental studies have shown, however, that there is ordering due to a negative mixing enthalpy of the transition metal cations. Therefore we attempt to understand the ordered solution NMC space and what intermediate phases exist. To do this, Metropolis Monte Carlo simulations we performed in order to calculate the change in Gibbs energy given by $G = U - TS + PV$. We ignore the PV terms since the volume changes are very small between end members and the mixed states. At every Monte Carlo step, a random swap of transition metals

or spin was chosen and the energy of the new and old states are then compared. The probability of accepting a swap is then given by

$$P_{1 \rightarrow 2} = \min[1, e^{\frac{-\Delta E}{k_B T}}] \quad (4.12)$$

The swap is accepted if this probability is greater than a random number between zero and one. For the ordered phases, the configurational entropy of cation orderings is ignored. The spin configurational entropy is assumed to be fully disordered as the simulation is performed well above the transition temperature for magnetic ordering in these materials [134]. To test this assumption, thermodynamic integration of the isochoric heat capacity was performed. From the fluctuation-dissipation theorem, the heat capacity can be computed as

$$\langle H^2 \rangle - \langle H \rangle^2 = k_B T^2 C_V \quad (4.13)$$

and from this, the entropy is given by

$$S = \int_0^T \frac{C_V dT'}{T'} \quad (4.14)$$

The prediction of entropy from this thermodynamic integration was within 10% of theoretical full spin entropy given by

$$S = N k_B [x_{Ni} + x_{Mn}] \ln(2) \quad (4.15)$$

Each Monte Carlo simulation started with a simulated annealing procedure to improve the convergence of the ordering. Starting from 1500K, the temperature was lowered by 150K until it reached 0K. Each temperature was run for 100 steps per lattice site. Then the temperature was raised to 298K, equilibrated for 100 steps per lattice site and a final sampling period of 100 steps per lattice site was performed. To simulate the possibility of a disordered phase that could be entropically stabilized at high temperatures, random swaps were always accepted and the energy was averaged over 5000 steps. The Gibbs energy per formula unit was then computed using the full configurational and spin entropy

$$G = \langle H \rangle - k_B T ([x_{Ni} + x_{Mn}] \ln(2) - x_{Ni} \ln x_{Ni} - x_{Mn} \ln x_{Mn} - x_{Co} \ln x_{Co}) \quad (4.16)$$

This procedure for the ordered and disordered phases was repeated for every composition in the fixed lattice size. Multiple lattice sizes were tested ranging from 5x5 to 10x10 to sample a range of orderings and compositions and a comparison of the energies for the ordered and disordered phases is seen in Fig. 4.6. The material is predicted to be disordered for nearly all compositions at 1700K, supporting the experimental ability to synthesize arbitrary compositions and quench to room temperature, locking in the disordered cations in the transition metal layer. At room

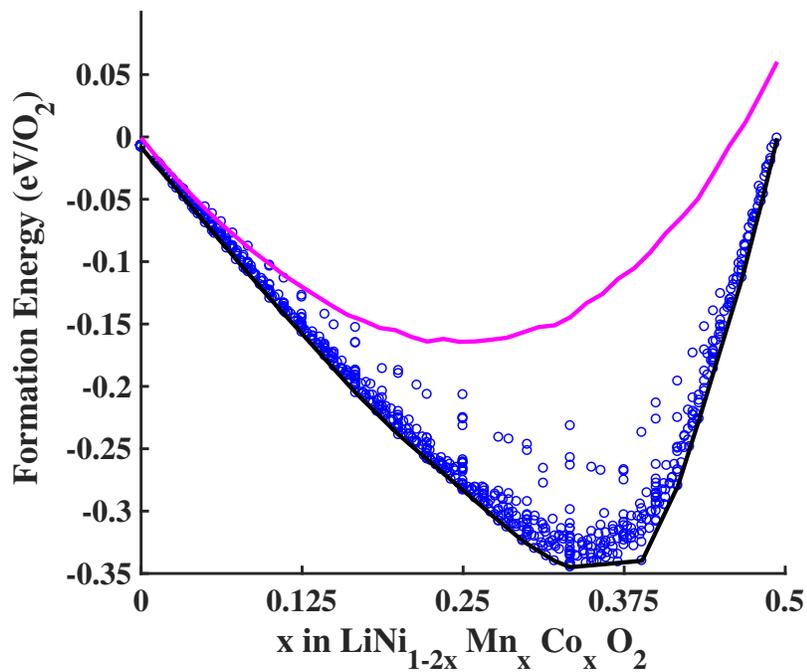


Figure 4.6: The prediction of energy as a function of composition for a one-dimensional slice of the total phase space. The ordered phases are given as blue dots and the disordered phases as a pink line. The large energy difference between the ordered and disordered phases demonstrates a large tendency for ordering. The multiple blue dots at a given composition are the energies of the ordered phases with different supercell sizes that were sampled. Reproduced from Ref [132], with the permission of IOP Publishing.

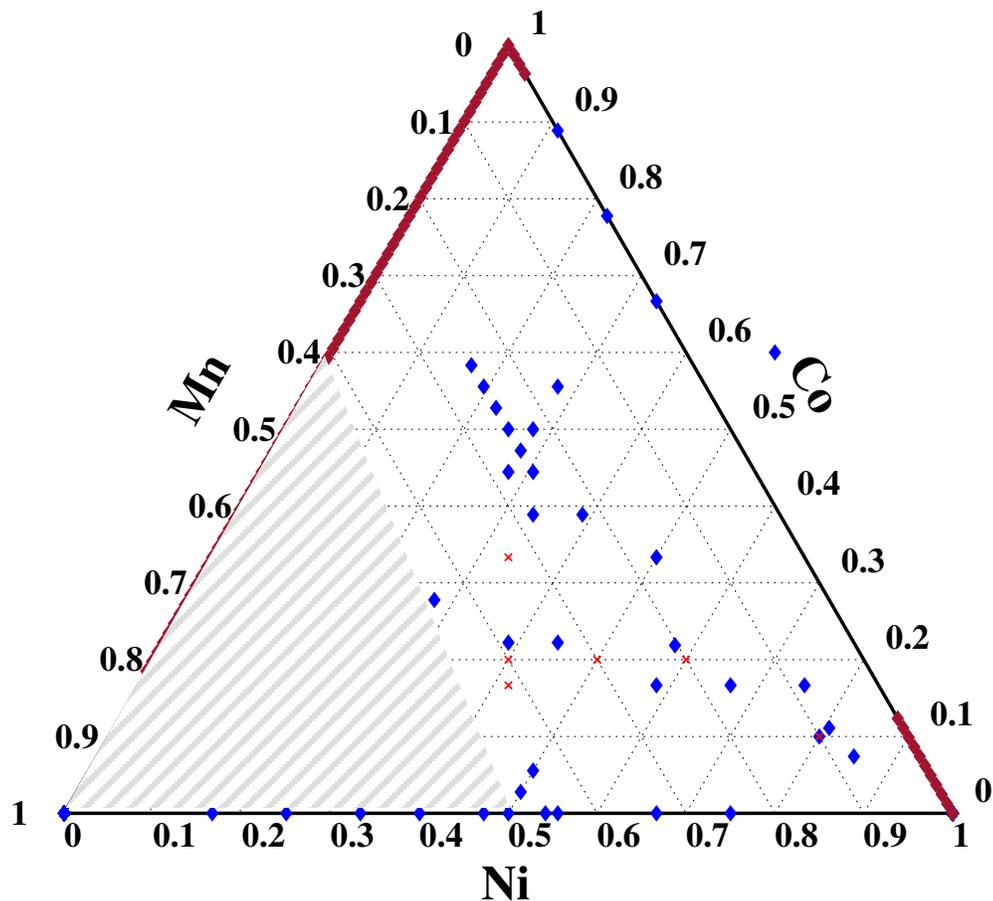


Figure 4.7: exp shown in red x from Ref. [135] disordered phases shown in dark red and ordered phases shown in blue. Computational data taken from Ref [132]. The high-Mn region determined to be unstable in the layered structure is blocked out.

temperature, the majority of the compositions are ordered however, as seen in Fig. 4.7. The predicted ordered phases, however, different from the common experimentally used compositions. Some well known experimental phases such as NMC6222 and NMC442 for example do not appear as predicted order phases but are near compositions that are predicted. This could suggest that these phases, which have nice proportions in synthesis, are used because they are close to the real stable phase and have convenient proportions for synthesis. On the other hand, given the error in DFT calculations, how confident are we that it is the experiment that is slightly off and not the prediction? In particular, NMC111 does not appear as a predicted order stable phase, but has been directly synthesized and measured to have at least local ordering [119].

To understand the sensitivity of the prediction on the exchange-correlation poten-

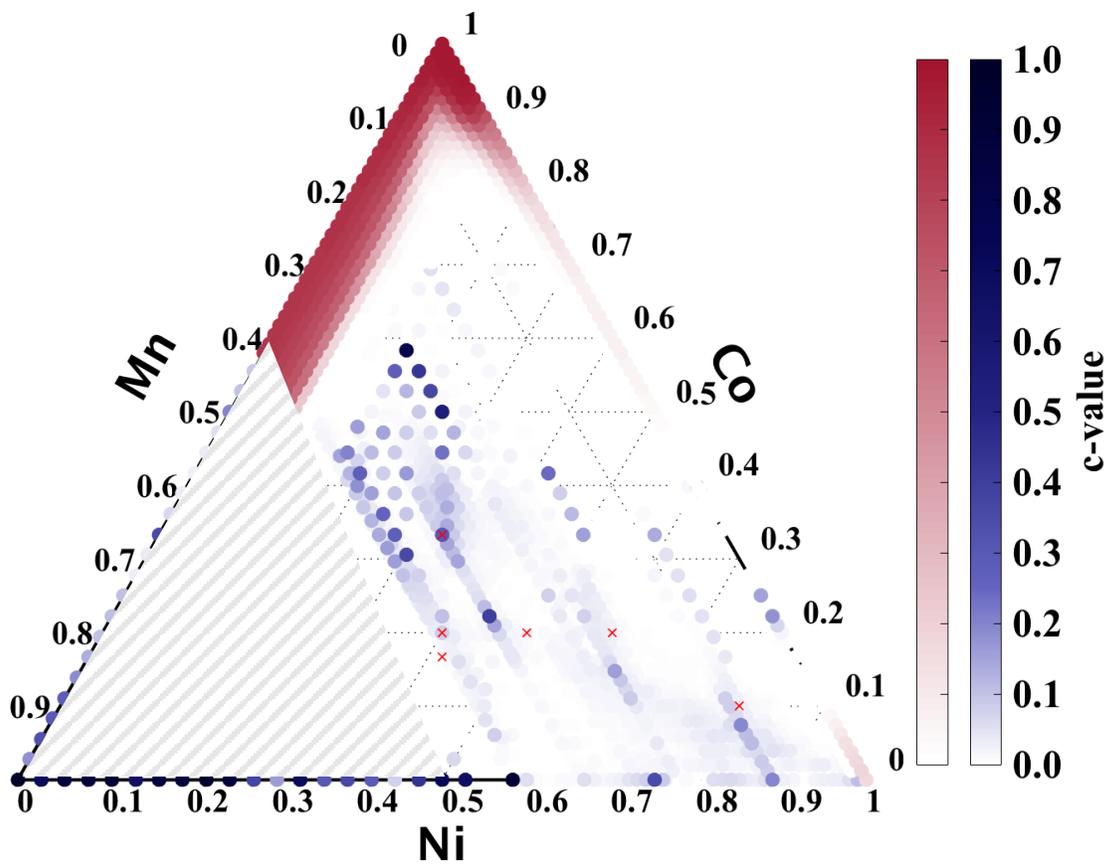


Figure 4.8: The prediction of ordered and disordered phases and the associated confidence value (c-value) are shown in blue and dark red respectively. Again high-Mn region is blocked out and the experimental phases from Ref. [135] are shown. [Reproduced from Ref [132], with the permission of IOP Publishing Copyright 2020]

tial approximation, we again utilized the BEEF-vdW ensemble of models. The model was fit for each of the 2000 functionals tested and the Monte Carlo simulations were repeated for both the ordered and disordered phases. For each functional, the convex hull was predicted. We employ a modified c -value for the prediction confidence of a phase being on the hull by counting the number of times each composition is predicted to occur and divide by the number of total hulls. The prediction of ordered phases and the region of disorder is now accompanied by a prediction confidence plotted in Fig 4.8. We can say very confidently that the majority of the phases are in ordered at room temperature as was predicted for the single functional. Not only do we recover the prediction of NMC11, but we also see a clustering of high confidence predictions for phases near but not exactly on the experimentally used phases, providing evidence for the possibility that the experimentally known phases are near but not perfectly on the most stable intermediate compositional phases. Some notable phases that are local peaks in the prediction confidence are NMC432 (c -value = 0.40) near the experimentally used phase NMC532, NMC22-3-2 (c -value=0.20) near the experimentally used NMC811 phase, and NMC 17-6-4 (c -value=0.16) near the experimentally used phase NMC622. The nominally low c -value of these phases is likely due to phases of the very similar compositions predicted as stable. If the confidence that some phase within a composition circle was computed, the prediction confidence would likely be much high as all of the prediction probability in that region would be counted.

Our final results suggest that given the complexity of the ternary phases search versus binary phase search were cluster expansion has been successful, and the small energy difference between two phases of very close composition, the prediction of stable orderings cannot be done with high confidence. These energy differences are too small for the uncertainty within DFT with respect to GGA functional choice and are worsened by the introduction of model uncertainty when training the cluster expansion and statistical noise from the Monte Carlo simulations.

We can, however, look at the type of ordering that is predicted in these materials by looking at the radial distribution functions (rdf) of the transition metals, which is the normalized occurrence of transition metals of the same species as a function of their distance. We plot the rdf for NMC111 and NMC811 using a supercell size of 18x18x12 and 20x20x12 formula units respectively in Fig. 4.9. When the rdf of each of the transition metals is compared to that of the all transition metal rdf, we can understand the tendency for mixing. In NMC111 seen in Fig. 4.9(a), there is a decrease in the prominence of like same species nearest-neighbor interactions (~ 2.8 Å), demonstrating that to first order, the transition metals will mix. For both materials, we see a tendency for next nearest neighbor separation of Mn atoms (~ 5.6 Å). In NMC811 seen in Fig. 4.9(b), we see that the Co-Co nearest neighbor interaction occurs at a higher rate than if the materials were disordered. The Ni atoms, on the other hand, appear to order similar to that of a disordered material suggesting that at such a high composition of Ni, the Ni atoms exist as a background for Co and Mn to mix within. In both cases, the ordering appears to vanish at larger distances of

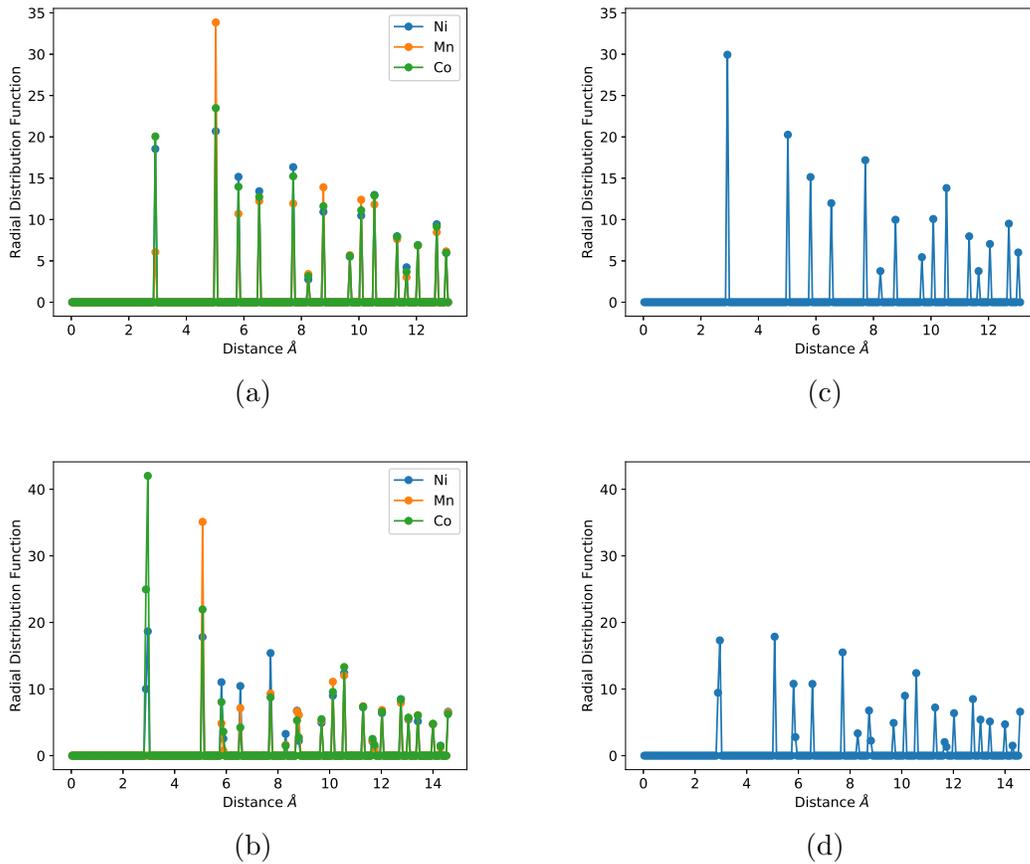


Figure 4.9: Radial Distribution Functions (rdf) for (a) NMC111 and (b) NMC811. For reference, the rdf for the all transition metals, which would represent a disordered phase in the supercells used are shown in (c) and (d). The comparison shows local ordering with less long ordering beyond 10 Å.

approximately 10 Å.

So far, we have limited ourselves to studying the fully lithiated NMC phases. To understand the ordering in the corresponding delithiated phase, $\text{Ni}_{1-x-y}\text{Mn}_x\text{Co}_y\text{O}_2$, this method was repeated. Using 72 DFT data points as training, the same model was fit and all phases were found to have a positive but small enthalpy of mixing. We assume that the stable orderings in the material are kinetically stabilized especially if the material is not cycled to a fully delithiated state. We, therefore, can assume that the $\Delta E_f \approx 0$ and energy of the fully delithiated NMC phase is the same as the energy of the mixture of the pure phases.

4.1.4 Prediction of Average Voltage

An immediate prediction we can get from the cluster expansion model is the average voltage for any phase of NMC. The average voltage is given from the change in Gibbs free energy of the intercalation reaction

$$V = \frac{-\Delta G_f}{e} \quad (4.17)$$

where the change in Gibbs energy of the reaction is

$$\Delta G = G_{\text{LiNMC}} - G_{\text{NMC}} - G_{\text{Li}} \quad (4.18)$$

Since we have predicted energy difference and assume the energy of the delithiated state is the same as the energy of a mixture of the pure compounds, we can modify Eq. 4.18 to include the experimental average voltages for each of the 3 pure materials to improve the accuracy of the prediction seen in Fig. 4.10.

$$V = \frac{-\Delta G_{\text{LiNMC}}}{F} + xV_{\text{Ni}} + yV_{\text{Mn}} + zV_{\text{Co}} \quad (4.19)$$

From this analysis, we are still left with a series of questions. We find that the average voltage is largely constant over the NMC range suggesting that there is little to no trade-off in this metric when decreasing Co content and increasing Ni content. But the average voltage does not entirely describe the operational voltage of the cathode. It is desirable to have a flat voltage curve as that provides a constant discharge potential over the majority of the battery’s capacity. However, a battery could have the same average voltage with a voltage profile that is much less constant. A prediction of the Li-vacancy ordering and energy at every state of Li content is therefore needed to compute the operational open-circuit voltage (OCV). This would require either a cluster expansion in the Li lattice for every NMC of interest or a very complex cluster expansion that can simultaneously account for transition metal ordering and Li-vacancy ordering. Unfortunately, both would require significantly more DFT data as training, and while possible, the amount of data required could be unfeasible.

Additionally, we find that at the fully lithiated state there is a strong tendency for ordering. At no lithiation, on the other hand, there is positive mixing enthalpy which could be partially responsible for the structural rearrangement seen at high states of charge. We want to understand the lithium content (and corresponding cell potential) at which the mixing becomes energetically unfavorable and leads to instability, and therefore what materials might handle higher states of charge.

Finally, the vibrational properties and the contribution to the Gibbs energy of these materials has been ignored. How much would these predictions change if the vibrational energy was included? The Debye model was shown to work well for LiCoO₂, but repeating for an extensive number of NMC phases would be expensive

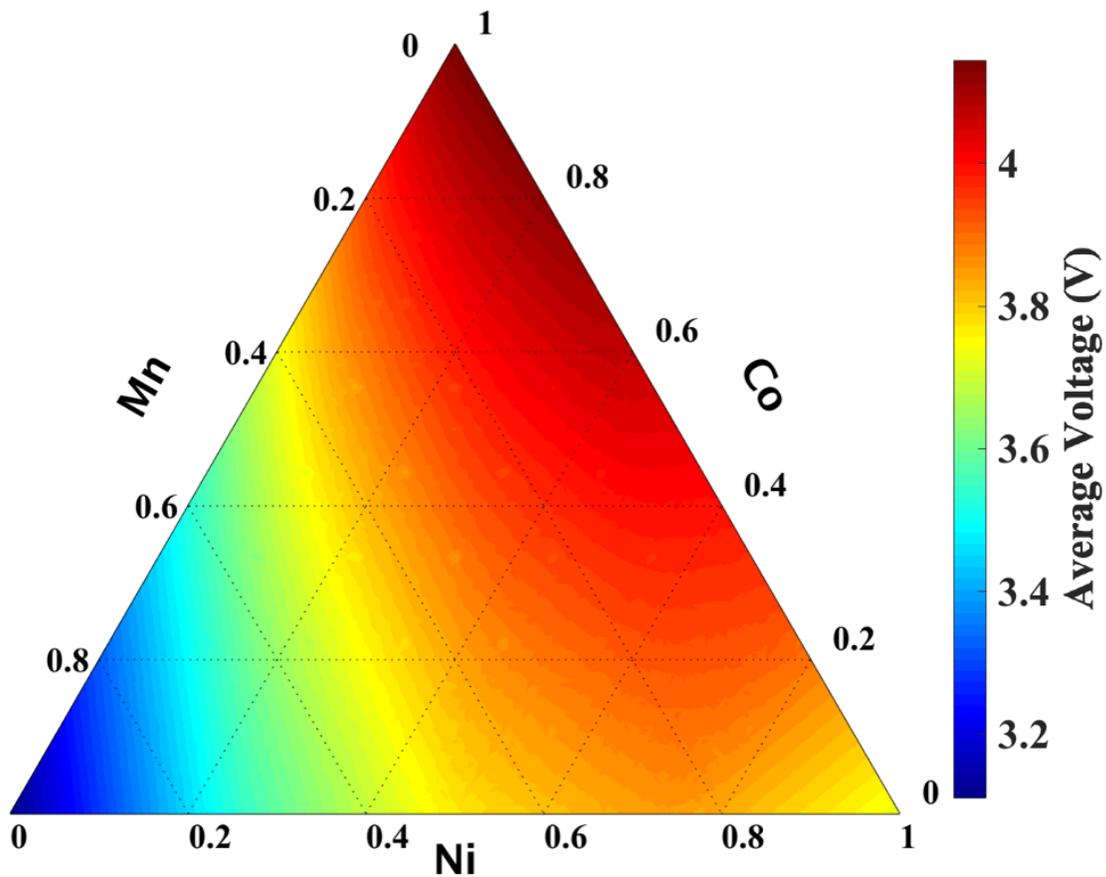


Figure 4.10: Predicted average voltage over all compositions of NMC. [Reproduced from Ref [132], with the permission of IOP Publishing, Copyright 2019]]

with DFT and impossible with cluster expansion, given that lack of atomic position dependence. All of the questions require a more capable model that is just as accurate or more than Cluster expansion but can account for the position of all atoms in an intelligent way.

4.2 Atom Centered Symmetry Functions

In this section, we utilize the atom centered symmetry functions suggested by Behler and Parrinello fed to an artificial neural network (BPNN) described in Chapter 2.2.1. As mentioned previously, these functions map the chemical environments of each atom to a single value that is invariant to rotations and translations of the material, as well as permutations of the atom labeling. The exact choice of optimal values for the hyper-parameters of these symmetry functions is in general unknown and picked by hand. To mediate the art of hyper-parameter selection, Bayesian optimization is employed for a fixed model complexity of two G_2 functions describing two-body interaction (Eq. 2.22) and two G_4 functions describing three-body interactions (Eq. 2.23). For each unique functional form, there will be an input node to the neural network for every possible two-body interaction for G_2 and every possible three-body interaction for G_4 , thus for the model complexity chosen, there will be 40 input nodes to each neural network to be trained. This model complexity was chosen to match the model complexity of the cluster expansion used previously and provides a good trade-off of computational efficiency with high accuracy as we will see later.

Not only does this machine learning potential provide improved accuracy, but it also provides increased versatility. The direct treatment of all species in the system allows for the prediction of lithium ordering beyond the transition metal ordering already discussed. The explicit position dependence of the symmetry functions and ultimately the predicted energy allows for the prediction of binding curves, equilibrium lattice constants, and forces. And the generality of the potential could allow for an efficient and accurate prediction of structural transitions as a function of lithium content as we will explore later. These structural predictions are available at speeds thousands of times faster than DFT and with an implementation much easier than constructing a cluster expansion for each structural phase. We will utilize this improved functionality to reproduce experimental results of lattice contraction and demonstrate a fast way to predict the open circuit voltage for an arbitrary NMC composition, something that has previously only been done on a composition to composition basis due to computational costs and with less accuracy than the results presented here.

To accomplish these tasks, we use all available data from the generation of the cluster expansion data set including the lattice constant optimization and the relaxation steps. This includes the fully lithiated LiNiMnCoO_2 in the O3 structure as well as the fully delithiated NiMnCoO_2 in the O1 structure. A collection of structures in the O3 phase with partial lithium content $0 < z < 1$ was also added. These new structures were generated in the same way as previously discussed where the

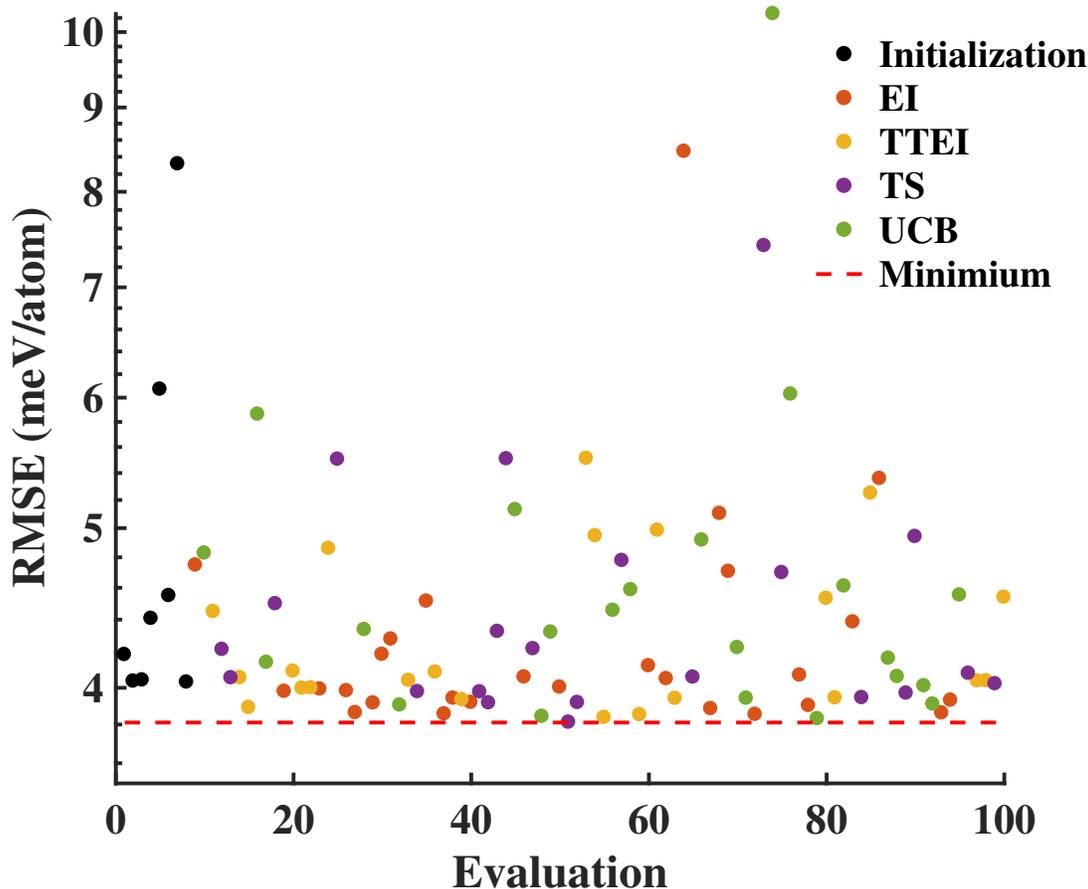


Figure 4.11: Dragonfly optimization. Reproduced from Ref [136], with the permission of AIP Publishing, Copyright 2019]

uniqueness of the structure was determined by the cluster interactions up to 6 Å. For the lithium vacancy structures, the interactions of Li were also added to the cluster interactions including transition metal-lithium interactions. Additionally, just as before, the lattice constants were optimized and the final atomic positions relaxed. All possible data from this process was kept. The data set of 12,962 points was then broken into a train, test, and validation set with an 80%, 10%, 10% split.

4.2.1 Hyperparameter Selection

The functional form of the atom centered symmetry functions described in Chapter 2 leaves a set of hyperparameters to be tuned for the best description of the system. One way to do this would be by way of guess and check, selecting a variety of guesses for the values, and comparing the errors of the predictions. Since we have chosen to

fix the model complexity, picking the model with the lowest RMSE is the same as using the Akaike information criterion that was previously employed. In this case, however, we have a continuous space of model parameters and cannot exhaustively search the whole space. A naive approach such a grid search would be expensive to sample a grid small enough to ensure optimum values. Instead, use the Bayesian optimization approach implemented in the Dragonfly package [137]. This approach begins by randomly sampling the phase space to begin to build a set of models for the value to be optimized as a function of the hyperparameters to search. It then utilizes a collection of acquisition functions that determine the next value of the function to evaluate. The Dragonfly implementation is unique in its combination of a collection of acquisition functions that employ various exploration versus exploitation tendencies. These acquisition functions include Gaussian process upper confidence bound (GP-UCB) [138, 139], Thompson sampling [140], expected improvement [141], and top-two expected improvement [142].

We optimized the hyper-parameters of the four unique functional forms of the symmetry functions used here. The resulting BPNN was then trained on the energy only of the training set to a stopping criterion of 1 meV/atom. The target function to optimize was then the energy RMS prediction accuracy of the trained BPNN on the test set. The results of this optimization are shown in Fig. 4.11 and give a root mean square error (RMSE) of 3.81 meV/atom on the test set. As the stopping criterion for each Bayesian optimization experiment was to train to 1 meV/atom RMSE on the training set, the model was likely overtrained slightly and therefore our search will pick a model with less tendency to overfit to the training data and provide the most generalizability to the test set.

We then use the best hyper-parameters from the energy only training to then select the force parameter that balances the training of forces and energy. This is done by training multiple models with force parameters ranging from 0.00001 to 0.01. For each model, the neural network was trained on the training set until it produced the best energy RMSE on the test set. The final model was then selected to have the lowest force RMSE from all of these models with similar energy RMSE of the energy only model. The best force parameter tested was $\alpha = 0.001$ and gave a final energy RMSE of 2.15 meV/atom and a force RMSE of 0.142 eV/Å on the test set seen in Fig. 4.12. To validate this hyper-parameter selection process which is likely biased to the test data set, we predicted the force and energy of the holdout validation set giving an RMSE 3.69 meV/atom and 0.129 eV/Å for energy and force respectively. We have finally arrived at a model with better RMSE error than the leave one out cross-validation of the cluster expansion, but that also includes Li-vacancies and atomic positions dependencies.

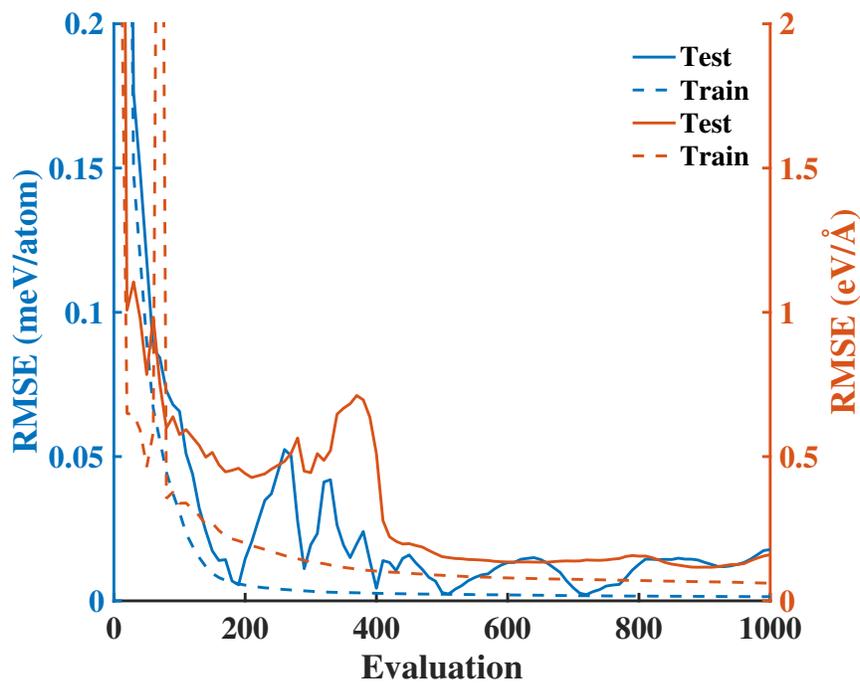


Figure 4.12: Plot of the force and energy RMSE on the test set as a function of number of epochs trained on the training set. Reproduced from Ref [136], with the permission of AIP Publishing, Copyright 2019]

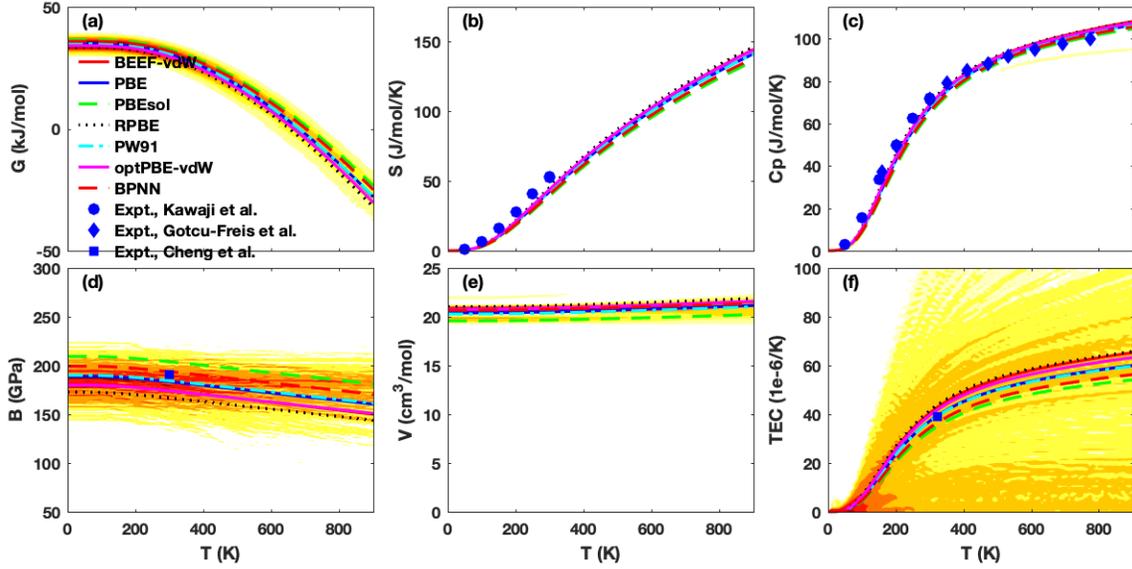


Figure 4.13: Debye-Grüneisen model prediction from the BPNN trained in this chapter compared to the full DFT predictions from Figure 3.5. These thermodynamic predictions include (a) Gibbs energy (b) entropy (c) isobaric heat capacity (d) bulk modulus (e) volume and (f) volumetric thermal expansion coefficient for LiCoO_2 .

4.2.2 Thermodynamic Properties

To test the accuracy of the machine learning potential trained, we feed the energy versus volume predictions to a Debye-Grüneisen model in the same procedure discussed in Chapter 3 for the case of LiCoO_2 . The same vibrational and elastic properties of Gibbs energy, entropy, bulk modulus, and thermal expansion were again predicted in the range of 0 to 900K. The comparison of the results versus the previous full DFT results is seen in Figure 4.13. Comparing to the DFT calculations over various functionals, we see that the BPNN does not deviate any more than any full DFT prediction deviates from the prediction of others. That is the added error of the BPNN in producing the DFT on which it was trained is not larger than the inherent uncertainty of the DFT prediction itself and for most properties is hard to visually pick out as it overlaps with all other predictions from DFT. We even see fair agreement of the BPNN in the prediction of the thermal expansion coefficient, which was previously discussed to have the largest tail and therefore the most sensitivity in the accuracy of the input energy predictions.

We next attempt to push to limits of the machine learning potential in an attempt to recreate the phonon spectrum in the harmonic approximation for LiCoO_2 . Within the harmonic approximation, the phonons are assumed to be non-interacting plane waves that are eigenvectors, \mathbf{e} , of the dynamical matrix, $\mathbf{D}(\mathbf{q})$, with wavevector \mathbf{q} .

$$\omega^2(\mathbf{q})\mathbf{e}(\mathbf{q}) = \mathbf{D}(\mathbf{q})\mathbf{e}(\mathbf{q}) \quad (4.20)$$

The dynamical matrix is defined in terms of the harmonic force constants determined by the second derivative of the energy with respect to the position of atom i in the α direction and atom j in the β direction.

$$\phi_{\alpha,\beta}^{i,j} = \frac{\partial^2 E}{\partial u_i^\alpha \partial u_j^\beta} \quad (4.21)$$

The dynamical matrix is then the mass reduced Fourier transform of these force constants, summing over all unit cell n .

$$D_{\alpha,\beta}^{i,j} = \frac{1}{\sqrt{M_i M_j}} \sum_n \phi_{\alpha,\beta}^{i,j} e^{-i\mathbf{q}\cdot\mathbf{r}_n} \quad (4.22)$$

The force coefficients are computed using finite difference as implemented in phonopy [143] using the energy predicted from the BPNN calculator. This showed improved results versus using the forces predicted analytically from the BPNN for one of the derivatives in the calculation of the force constants. The predicted spectrum, shown in Fig. 4.14, shows rather unrealistic looking dispersion curves as the prediction required accurate computation of the second derivative of energy. The neural network on the other hand was only trained on the first derivative of energy and energy itself leading to a large error in the force constants. This is in contrast to the results of the Debye model which only requires smooth binding curves and therefore gives much better results.

The poor recreation of the phonon spectrum is likely due to two main factors. Firstly, given a perfect calculator, the occurrence of imaginary frequencies would indicate that the material structure is not dynamically stable. Therefore, the machine learning potential may incorrectly believe that the structure would be more stable if it were to deviate from the experimentally seen O3 structure. The addition of DFT data with LiCoO_2 in some other metastable structures may aid in the calculator learning the absolute stability of the O3 structure.

The second possible source of error is due to another deficiency in the dataset. The calculation of the force constants requires accurate prediction of the energy change due to moving a single atom small distances away from equilibrium while keeping the positions of the other atoms fixed. Not only does the absolute energy of this prediction have to be accurate, but the curvature of the resulting energy curve must be correct. Thus if the data set includes only structures that are very close to equilibrium, these slight atomic movements could fall out of the region of highly accurate predictions. In fact, a previous study using the performing finite-difference on the energy instead of forces to find the force coefficients found that to reach convergence on the resulting force coefficients, the input DFT had to be converged to a level of 10^{-10} eV/atom [61], orders of magnitude more precise than the input data used here and the accuracy

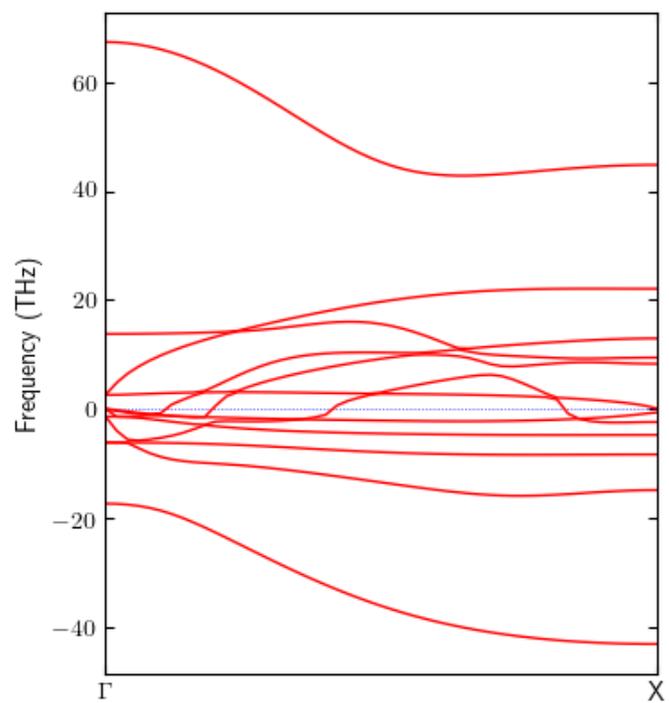


Figure 4.14: Computed phonon spectrum for LiCoO₂ using the BPNN calculator.

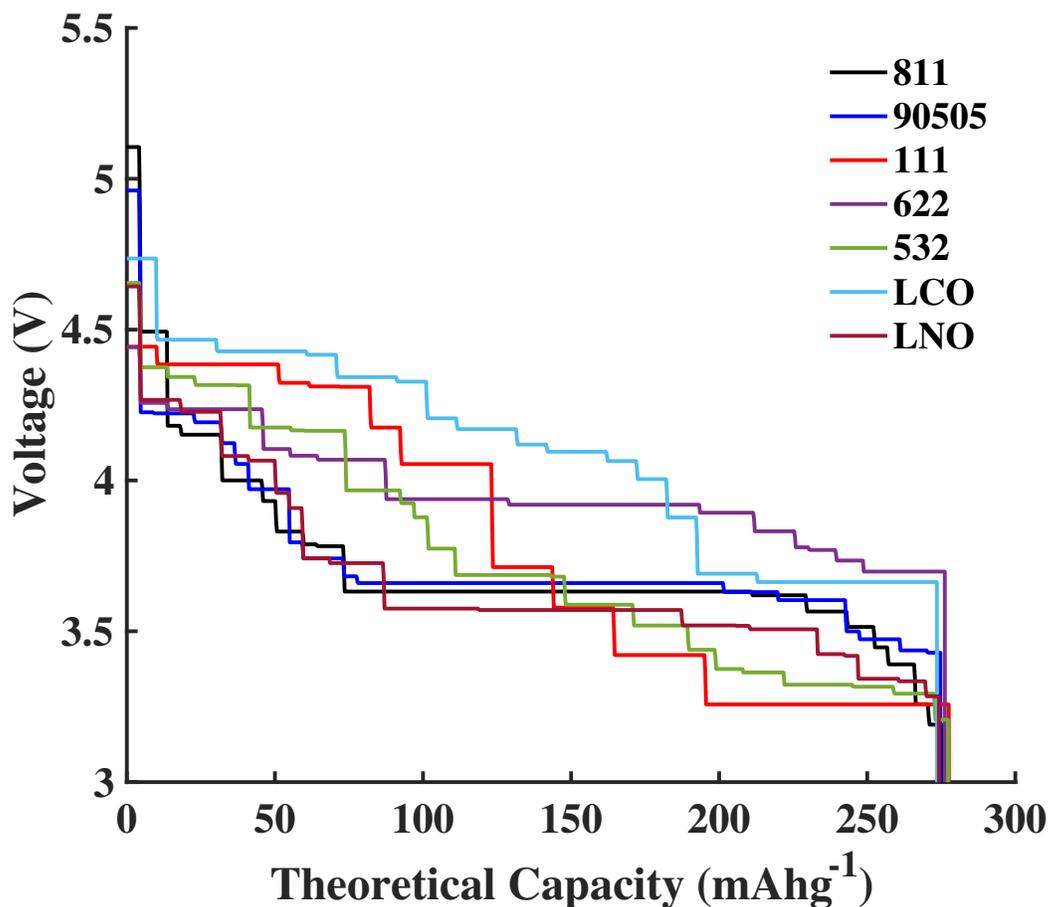


Figure 4.15: Predicted open circuit voltages for a series of NMC compounds.[Reproduced from Ref [136], with the permission of AIP Publishing, Copyright 2019]

of the BPNN model. Thus to improve the predictions in this aspect, the addition of DFT data from some sample finite-difference phonon calculations for a collection of NMC phases could be included.

In summary, we see that the prediction of phonons requires a level accuracy and interpolation in the phase space far beyond what the current BPNN model is capable of. This is in contrast to using the Debye model which can be reproduced not only more quickly but with a level of accuracy effectively at the level of the underlying DFT data.

4.2.3 Open Circuit Voltage

The most important quantity of a battery cathode is the open-circuit voltage at every point in the full capacity of the electrode. This determines the total energy of the battery which is equal to the area under the voltage-capacity curve. The shape of the curve also determines how desirable the material is as a cathode. While a material may provide a high potential at high states of charge, what is more important in real applications is a constant, flat potential.

To measure the open-circuit voltage of a cell experimentally, it is first charged fully to a set cutoff voltage. The cell is usually cycled a few times to get an estimate of the total capacity of the cell. For the high fidelity experiment, the cell is then discharged at a fixed current which is conventionally expressed in terms of a C-rate. This C-rate provides a normalized measure of the current draw with respect to the total capacity of the cell. The C-rate is the current of the discharge represented as the fraction of the current rate in a one hour discharge of the battery. Therefore a discharge rate of 1C, will discharge the battery in one hour. If the battery is discharged in 10 hours, then it is only using 1/10 of the current that is used for a one hour discharge and therefore has a rate of C/10. As this discharge represents a non-equilibrium process, the experimental rate should be very low and high fidelity open circuit voltage curves are commonly run over a multiple-day discharge. During this constant rate discharge, the potential of the cell is constantly measured and plotted with respect to the total amount of charge passed (the current multiplied by time elapsed represented as mAh). As the total capacity passed in the experiment is dependent on the amount of active electrode material, the capacity is divided by the mass of the discharged active material and represented as gravimetric capacity density (mAh/g). The experiment is finished when the bottom cutoff voltage is reached and all of the capacity is used.

As this experiment can require days to complete for a single run, we provide a fast computational alternative for the prediction of the open-circuit voltage curve. To find the open-circuit voltage we employ Grand Canonical Monte Carlo simulations where the chemical potential of Li rather than the number of Li atoms in the system is fixed. Therefore, we can insert and remove lithium atoms depending on the following acceptance criteria s given by [144]

$$P_{insert} = \min \left[1, \frac{V}{\Lambda^3(N+1)} \exp(-\beta[E_2 - E_1 - \mu_{Li}]) \right],$$

$$P_{remove} = \min \left[1, \frac{N\Lambda^3}{V} \exp(-\beta[E_2 - E_1 + \mu_{Li}]) \right],$$

where Λ is the thermal de Broglie wavelength

$$\Lambda = \frac{h}{\sqrt{2\pi mk_B T}}, \quad (4.23)$$

N is the number of lithium already in the material, and V is the volume accessible to the lithium given by [145] $V = V_{cell} - N \frac{4}{3} \pi r_{Li}^3$. Here we use the van der Waals radius of lithium $r_{Li} = 182 pm$. [146] After every trial of addition or removal, a set of canonical Monte Carlo steps are performed equal to the number of total possible Li sites in the lattice to aid in finding the best possible arrangements of lithium ions.

This Grand Canonical Monte Carlo can be used in two ways. The first is to simulate the discharge of a battery. In this case, the chemical potential is set to that of lithium in standard crystalline fcc state equal to the Gibbs energy per atom. For the prediction of the Li free energy, especially when compared to the free energy of lithium comparing compounds, it has been previously demonstrated that referencing to a lithium-containing material with known formation energy can improve accuracy as the error in DFT formation energy are correlated [62]. For this purpose we use LiCl with a measured formation energy of $\Delta H_f^{exp} = -4.231 eV$ [147]. We then compute the energy of Cl_2 and LiCl and find the energy of Li from

$$E_{Li} = E_{LiCl} - \frac{1}{2} E_{Cl_2} - \Delta H_f^{exp} \quad (4.24)$$

In this simulation, the lithium will continue to enter into the lattice until all of the spots are filled and the lattice can no longer be filled. This method is particularly useful when the final amount of lithium is not known as Li can be added and the positions relaxed until the chemical potential of Li in the cathode matches that of elemental lithium and no new moves additions of Li are accepted. For computational ease, although we know the total number of Lithium sites, we use this method to quickly simulated the voltage curve for a large range of NMC compositions. This is done by saving the composition and the energy after the end of the canonical Monte Carlo steps. From this collection of compositions and energies, a convex hull can then be constructed and the voltage can be calculated as

$$V = - \frac{(G_{Li_{x_2}MO_2} - G_{Li_{x_1}MO_2} - (x_2 - x_1)G_{Li(s)})}{e(x_2 - x_1)} \quad (4.25)$$

The second method involves sweeping over the chemical potential of lithium by varying the applied potential, U , and expressing the new chemical potential of lithium as

$$\mu_{Li} = \mu_{Li}^0 - eU \quad (4.26)$$

The Gibbs energy of the system can be found from integration as

$$\mu_{Li} = - \frac{\partial G}{\partial x_{Li}}. \quad (4.27)$$

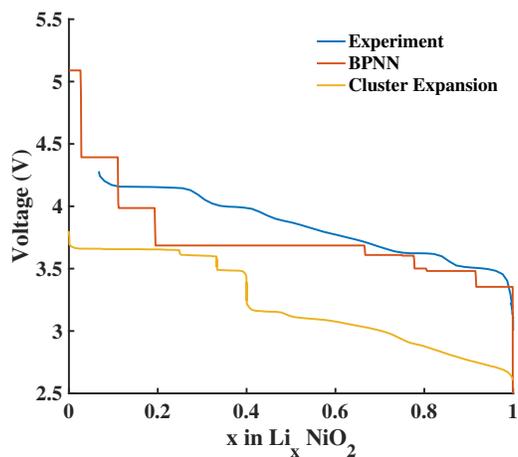
We can then find the convex hull and compute the voltage as before. The voltage can also be computed directly with respect to the Li/Li^+ reference from the applied

potential as $U(\text{Li}/\text{Li}^+) = -\frac{\mu_{\text{Li}}}{q_e}$. The average number of lithium present sampled over the simulation at that chemical potential would then represent the lithium content that corresponds to that open-circuit cell voltage at equilibrium.

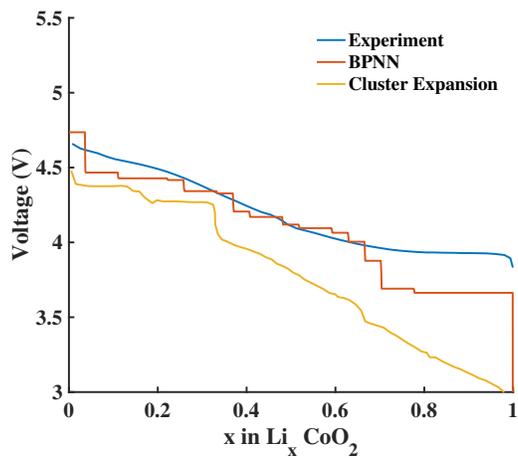
This first method was performed for a collection of NMC phases in Fig. 4.15. For each $\text{LiNi}_{1-x-y}\text{Mn}_x\text{Co}_y\text{O}_2$, composition, the transition metal orderings were determined using the cluster expansion trained previously using Canonical Monte Carlo as explained above for a supercell that was 5x4 conventional cells large for a total of 240 atoms per supercell. Throughout the Grand Canonical simulation, the transition metal ordering was fixed. The lattice constants of the fully lithiated material were optimized using an equation of state fit first to the expansion in the a=b lattice directions and then for the c-lattice constant. The atomic positions were then relaxed to a max force of 0.25 eV/Å or 10 relaxation steps. To ensure the minimum lattice constants were found, the process of lattice optimization and internal atomic position relaxation were performed. This second step of relaxation improved in finding a reasonable low force configuration compared to always relaxing to the force cutoff and not limiting to 10 relaxation steps for the first relaxation. All of the lithium atoms are then removed to simulate the fully charged state. The lattice is then optimized again twice, this time to a max force of 0.5 eV/Å. Additionally, to simulate the lattice contraction and expansion during cycling, the lattice constants are re-optimized after every Grand Canonical Monte Carlo step. The internal positions are however not relaxed as the positions of the oxygen atoms would change depending on the current positions of the Li atoms and therefore create preferential spots for Li to sit.

We find from our result of open-circuit voltage for a collection of NMC materials that while the prediction of average voltage did not vary largely, the voltage profiles vary significantly more. We see the biggest slope for the NMC111 compositions with a continuous decrease in slope as the composition moves along the $x_{\text{Co}} = x_{\text{Mn}}$ line of the phase space. We also predict different nominal voltages, which is the voltage at which the majority of the capacity delivers the most constant potentials. We find that LiCoO_2 has the highest nominal voltage as expected, with trends in decreasing nominal voltage that cannot be quickly explained. For example, LiNiO_2 is predicted to have one of the lowest nominal voltages as do other high Ni content materials such as NMC811 and NMC90505. But the Co content itself is not the sole descriptor of nominal voltage. While it does increase in general with increasing Co content, NMC622 is predicted to have a higher nominal voltage compared to NMC532 with the same amount of Co. Additionally NMC811 and NMC90505 are predicted to have almost the same nominal voltage and therefore no trade-off in that aspect when changing Co content.

To benchmark these results, we compare the prediction from the trained Behler-Parrinello neural network (BPNN) to that of experimentally measured voltage and previous computational work. This is done for both LiCoO_2 and LiNiO_2 in Figure 4.16a. The previous computational works [148, 124] both employed a cluster expansion for the Li-vacancy ordering much like the method described above. A set of DFT



(a)



(b)

Figure 4.16: Prediction of open circuit voltage from the BPNN for (a) LiNiO_2 compared with recent experiment from Ref. [110] and previous predictions using Cluster Expansion trained on DFT in Ref. [148], and (b) for LiCoO_2 compared with experiment from Ref. [149] and previous predictions using Cluster Expansion trained on DFT in Ref. [124].

calculations were used to train the model and Grand Canonical Monte Carlo simulations were performed over a sweep of chemical potentials. We find that although our supercells were much smaller than those used in the cluster expansion due to the increased computational cost of our method, we see a better agreement with experiment. In particular, the cluster expansion for LiCoO_2 shows a constantly decreasing voltage near $x_{Li} = 1$ while both the experiment and the BPNN give a plateau. For the case of LiNiO_2 , the machine learning potential is unable to correctly capture the plateau and then drop at around $x_{Li} = 0.3$. This flat portion of the curve is associated with a structural phase transition from the O3 layered ordering to an H1-3 stacking that will be discussed in more detail later. While the energy difference in this phase transition that also occurs in LiCoO_2 , was small enough to ignore, it appears to be crucial in the case of LiNiO_2 . The cluster expansion employed separate models for each of the structural phases of LiNiO_2 and therefore has a slight improvement in the qualitative shape, although does not perfectly recreate the voltage profile.

In the next section, we explore in more detail the structural phase transitions that occur in LiCoO_2 and LiNiO_2 during cycling and test the ability of the machine learning potential to extrapolate to these other structures which have not been explicitly included in the data. We also use the machine learning potentials ability to predict the layer separation to probe the lattice contractions that occur at high states of charge and that have been linked to capacity fading.

4.2.4 Structural Changes During Cycling

The two materials of LiCoO_2 and LiNiO_2 share similar yet slightly differing structural evolution during cycling. As mentioned previously they both share the O3 layered structure in the fully discharged state ($x_{Li} = 1$). As each material is charged and Li is removed, they both undergo a transition to monoclinic symmetry. With no distortions in the monoclinic cell that breaks the $R\bar{3}m$ symmetry, these two lattices are equivalent as seen in Fig. 4.17. In the case of LiNiO_2 , the in-plane atomic spacings in the a and b-lattice directions are distorted slightly. More specifically, the a_{mon} lattice constant is stretched larger than its corresponding value that would regenerate the O3 structure. In LiCoO_2 , the monoclinic c-lattice spacing is stretched. The c-lattice vector for monoclinic cells is given by the vector connecting a lithium atom to the nearest Li atom in the layer above and is therefore tilted with respect to the O3 c-lattice constant. The stretching in the monoclinic cell then leads to shearing in the layers and therefore would cause more mechanical stress than the monoclinic distortions of LiNiO_2 . As the lithium content of these two materials is completely depleted, they will transform into the O1 rhombohedral stacking in which the transition metals are aligned vertically and there are no shifts in the layers. In the intermediate at around x_{Li} , a hybrid rhombohedral phase exists known as H1-3. In this structure, the layers are aligned for one repeated layer and then shifted for another to create a layering patterning that is six layers large seen in Fig. 4.17(a).

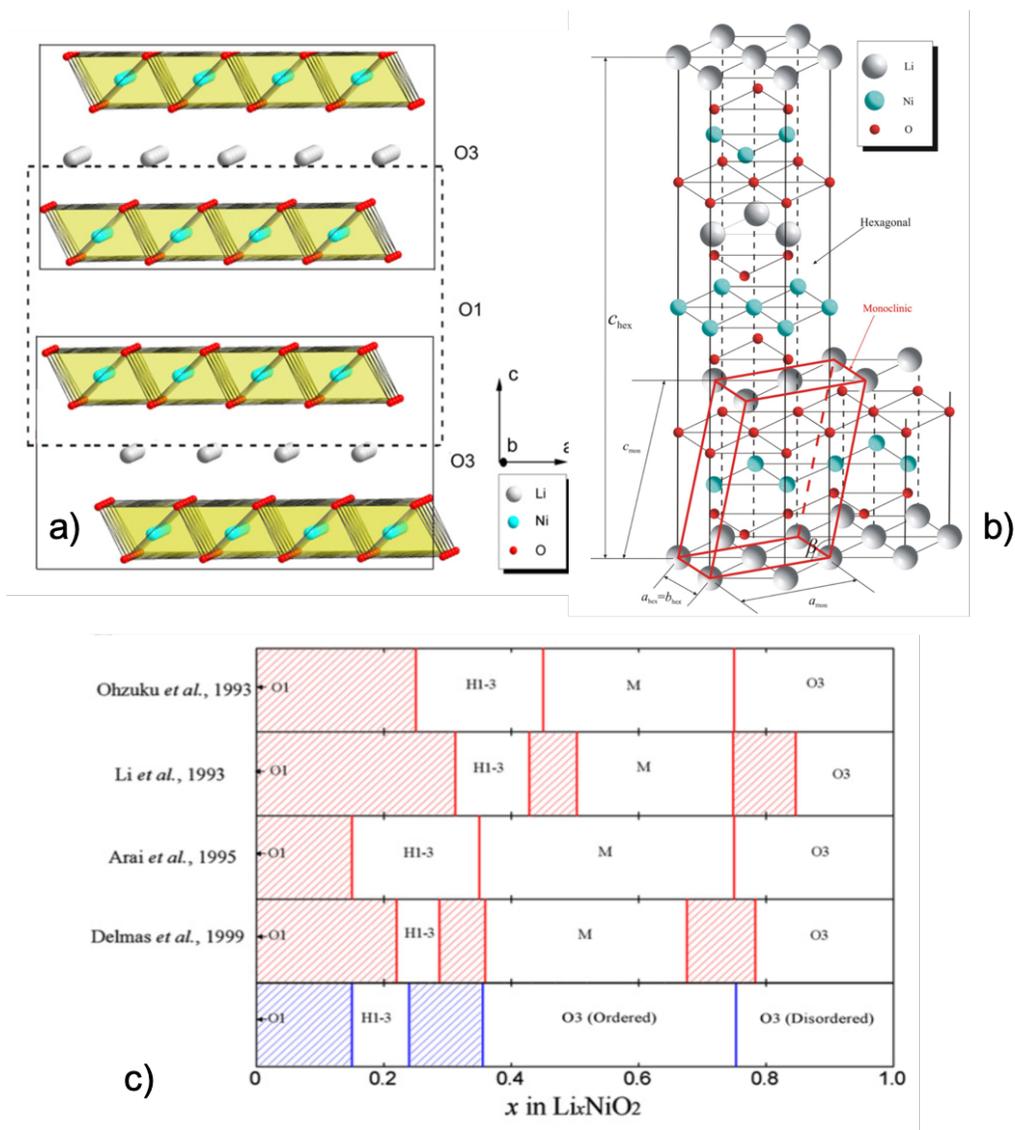


Figure 4.17: The various structures of $LiNiO_2$. The relation between the H1-1, O3, and O1 layer stackings are seen in (a). The whole picture shows the H1-3 pattern, while the dotted line shows the O1 stacking, and the solid line the O3 stacking. The correspondence between the hexagonal and monoclinic symmetries is seen in (b). And the phase diagram as measured from a collection of experiments as well as the predictions in blue from the computational phase modeling studying of Chang *et al.* The phase diagram for $LiCo_2$ is largely similar. [Reprinted from Ref. [150]. Copyright 2012, with permission from Elsevier]

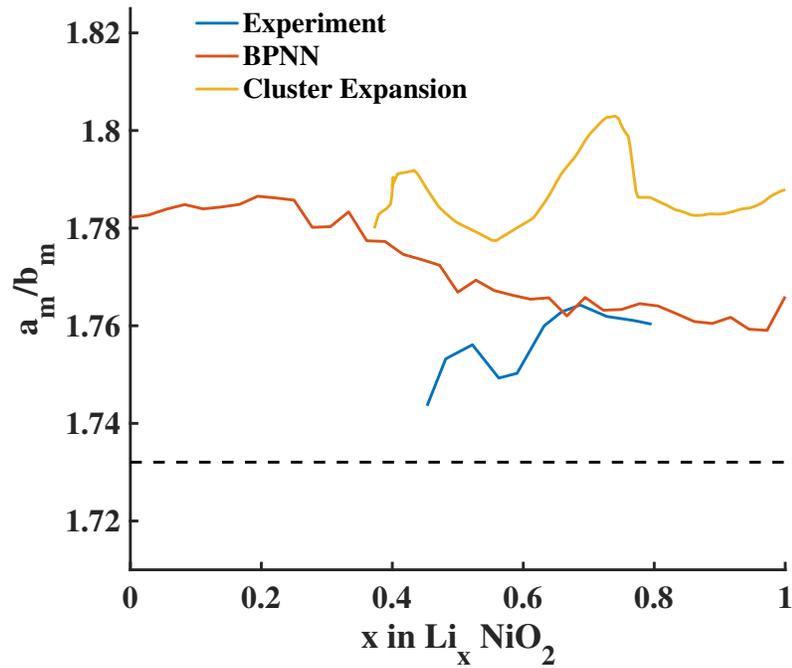


Figure 4.18: Test of the prediction of monoclinic distortions in LiNiO_2 . For comparison, experimental predictions from Li *et al* [109] and cluster expansion trained on the a_m/b_m ratio by Arroyo y de Dompablo *et al.* [151]. We also show the line corresponding to a ratio of $\sqrt{3}$ which would indicate no distortion from the hexagonal phase.

Monoclinic Distortions in LiNiO₂

We first look at the predictions of monoclinic distortions in LiNiO₂. The origin expansion of the b_m lattice constant has been understood through Jahn-Teller distortions in the transition metal octohedra. [151]. Previous first-principles calculations by Arroyo y de Dompablo *et al.* fit a cluster expansion to the ration of a_m/b_m in an attempt to understand the evolution of the monoclinic distortion as a function of lithium compositions. We thus use this distortion to test the ability of our machine learning potential to extrapolate to similar yet differing structural phase spaces than the one it was trained on. A series of Monte Carlo simulations were performed for various Li compositions in a 144 atom supercell. At each lithium compositions the all lattice constants were allowed to relax independently. We then plot the a_m/b_m for our predictions as well as the previous cluster expansion and experimental observations. A material with no Jahn-Teller distortions representing the O3 structure would have an $a_m/b_m = \sqrt{3}$.

We see that we match the nominal degree of distortion closely with respect to the experiment but do not capture the qualitative shape. Both the experiment and the cluster expansion show a maximum in the distortion at around $x_{Li} = 0.5$ and 0.7 . In the work by Arroyo y de Dompablo *et al.*, this was explained by two lithium-vacancy ordered phases that maximized the Jahn-Teller distortion.

Lattice Contractions at High States of Charge

As the concentration of Ni increases, although the capacity of the cathode increases, there is a large capacity fade over many cycles. The capacity retention is worsened by cycling the battery to high cutoff voltages above 4.2V. The mechanism of degradation has been linked to the phase transition from the H1-3 phase to the O1 phase that is accompanied by a rapid contraction in the c-lattice layer separation [152]. This lattice contraction increases in general with increasing Ni content and therefore the capacity fade will worsen with increased Ni. Another hypothesis is that the surface reactivity at high states of charge and therefore low Li content. A study of NMC811 proposed that as the lattice contraction from the structural phase transition was reversible, as demonstrated by comparing refinement data from pristine and cycled NMC811, that there was no cracking and this is not the explanation for capacity fade [153]. They saw improved capacity retention when modifying the electrolyte and therefore suggested that the electrolyte degrades the highly reactive low Li surface. These competing ideas were unified in a study over a range of high Ni cathode materials by Ryu *et al.* [3]. They demonstrated that the lattice contraction was the cause of the increased reactivity due to the creation of cracks that propagate to the surface and create channels for the electrolyte to penetrate and react. We attempt to test the trend of increasing Ni content on the degree of contraction at high states of charge. To do this, the H1-3 and O1 phases were simulated with Canonical Monte Carlo to find the most stable Li ordering. The lattice constants were then optimized and

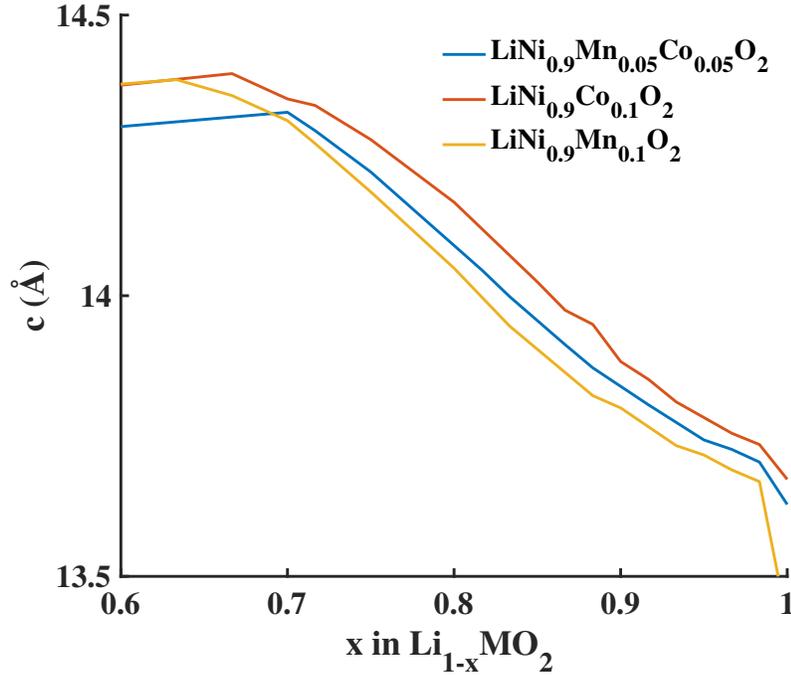


Figure 4.19: Prediction of the lattice contractions that occur during the high states of charge for three NMC compositions all with 90% Ni content.

the internal coordinated fully relaxed with the machine learning potential. A convex hull of these phases was fit and the c -lattice separation for every point on the hull is plotted in Fig. 4.19

From this, we predict the contractions are not strictly due to the total amount of Ni content, but depend secondarily on the composition of Ni and Mn, with equal proportions of the two given the smallest percent change in the c -lattice constant. We calculate that the change in lattice constant

$$\Delta c = \frac{c_{min} - c_{max}}{c_{max}} \quad (4.28)$$

is the largest for $\text{LiNi}_{0.9}\text{Mn}_{0.1}\text{O}_2$ with $\Delta c = -6.7\%$ and the smallest for $\text{LiNi}_{0.9}\text{Mn}_{0.05}\text{Co}_{0.05}\text{O}_2$ with $\Delta c = -4.8\%$ which compares to an experimentally seen contraction for this same composition of $\Delta c = -5.6\%$ [3]. Overall we find that the high Ni content still dominates the lattice dynamics and there is only a small secondary effect from the relative composition of Mn and Co.

4.3 Summary

This chapter was about developing efficient computational tools that enable the understanding of the ordering, structure, performance, and dynamics of layered transition metal cathode materials. We first attempted to understand the transition metal ordering for the fully discharged phase $\text{LiNi}_{1-x-y}\text{Mn}_x\text{Co}_y\text{O}_2$ with the use of cluster expansion trained on DFT data. Model selection techniques were used to select the best model that balanced accuracy with simplicity. The ordering of the transition metals was predicted through Monte Carlo simulations using the cluster expansion as the energy calculator. The resulting convex hull which included simulations of disordered phases with full configurational entropy was not in complete agreement with the experimentally studied phases. To understand the sensitivity of this prediction on the errors from DFT and to provide a quantitative metric of confidence for the predicted stable phases, an ensemble approach was used. The cluster expansion was retrained for every DFT functional in the ensemble to create an ensemble of models that each generate a convex hull from Monte Carlo simulation. From this, a prediction confidence for each phase is computed as the fraction of hulls in the ensemble that predicts the phase to be stable. From this final diagram for the prediction of phases, we see clustering of probability distribution centered around peaks in the probability that are very near experimentally seen phase. This leads to the prediction that a phase near that composition is stable but there is relative uncertainty for the exact composition of that phase and therefore conclusive predictions of stable phases are difficult given the combined error of DFT, cluster expansion, and Monte Carlo simulations.

Next, we trained a machine learning potential on the entire pseudo-quaternary phase space where both the transition metal compositions and lithium composition were varied within assuming the same layered structure. The hyperparameters were optimized using a Bayesian optimization search. We then demonstrated the potential’s ability to recreate the thermodynamic properties of LiCoO_2 using the Debye model and compared it to the previous prediction from Chapter 3. We found a negligible added error in terms of thermodynamic properties as this material is well within the phase space the machine learning potential was trained on. This result is in contrast to the predicted phonon spectrum within the harmonic approximation. The calculation of the force constants through the finite-difference of the energy for atomic displacements lead to imaginary frequencies suggesting either the model was wrongly mischaracterizing LiCoO_2 as dynamically unstable, or deficiencies in the data which largely only samples near equilibrium. We employed an efficient method using Monte Carlo to predict the open circuit voltage for a variety of NMC compositions. This method was benchmarked with respect to experiment and other computational studies showing good agreement with experiment versus the compared computational work. We then tested the machine learning potential’s ability to describe the phase transitions and structural dynamics during cycling. Despite not being trained on the

monoclinic structure, the nominal value of experimental distortion was reproduced for LiNiO_2 . We also predict the lattice distortion of Ni-rich phases to understand the secondary effect of composition on the lattice contractions linked to degradation. In all, we show the capabilities and limitations of the machine learning potential used. The biggest limitation in the prediction of the phonon spectrum, however, can be avoided through the use of the Debye model, and therefore the machine learning potential provides a large step towards accurate computational predictions of these cathode materials at speeds much faster than Density Functional Theory.

Chapter 5

Advanced Machine Learning Potentials

The goal of machine learning potentials is to map the identity and coordinates of a material to some feature space that describes the local environment of each atom in the material. This featurization can then be fed to a neural network that can be trained *ab initio* energy and force data for the material as we saw in the previous chapter. The ultimate goal is to achieve computational evaluation costs similar to classical molecular dynamics, with accuracy approaching that of the underlying quantum mechanical *ab initio* data on which the machine learning potential was trained. When training these models, the underlying symmetries of the physics and of the material must be preserved, including the permutation of atom labels, rotation, and translation of the entire material. The first step in realizing a machine learning architecture that obeys these symmetries is to develop featurizations of the material that are invariant to these symmetries and transformations. Previous architectures utilizing symmetry invariant descriptors are have shown to be too restrictive and may lose information and mischaracterize two truly distinct atomic environments as similar. Additionally, the selection of hyperparameters within these symmetry preserving features remain challenging to perform in a rigorous, deductive manner.

A more efficient method would entail an algorithm that can learn the featurization of materials by having an architecture that can describe bonds, bond angles, and long-range interactions all while preserving the known symmetries of the material and physics. In this chapter we briefly discuss field of machine learning potentials that are both fingerprint based and deep learning based, in which the architecture itself learns the important features, methods.

5.1 Review of Current Methods

In the previous chapter, we explored two methods to generate a model for the energy of a material. The first was a shallow learning method in cluster expansion,

and later atom centered symmetry functions were used. In both of these cases, the material was first fingerprinted in some way to generate descriptors for the energy. These descriptors were then passed to a regressor: multivariate linear regression, or a feed-forward neural network. A series of fingerprint-based methods exist to generate machine learning potentials. Based on the original atom centered symmetry functions passed to neural networks used within this work [32], modified Behler-Parrinello symmetry functions have also been used. [154, 155] Additionally, as we saw previously, the method to select the hyperparameters corresponding to the symmetry function could not be performed in a systematic way. Another approach involved expanding the neighborhood in terms of spherical harmonics and using a bispectral analysis of the neighborhood as the descriptor [156]. These bispectral components have also been used with linear regression in a method known as a spectral neighbor analysis potential (SNAP) [157]. The relatively large sensitivity of these spectral components on the slight changes in the atomic environments lead to the development of the smooth overlap of atomic potentials (SOAP) method where the density of the atoms are not treated as delta-functions as done in SNAP, but rather smooth Gaussian functions. [158]

All of these methods still require the tuning and selection of many hyperparameters or expansion cutoffs. Some suggest well-defined methods to define unique descriptors for a material. More simplistic and computationally efficient descriptors have also been suggested. For example, the Coulomb matrix uses Coulomb like function depending on distance and nuclear charge. [159] The overlap matrix uses the covalent radius of each element to defines a sphere around every atom in a material and calculates the overlap of these sphere as the descriptor of the system [160] Another method generates a graph based on the atomic identities and bonds that are defined by the covalent radii and Voronoi surfaces of each atom. [161]

Other state-of-the-art methods are deep neural networks that learn underlying representation automatically but in a more complex way. These include method such as deepMD [162] and deep tensor [163]. Some attempt to use the atomic gradient instead of total energy in a method known as gradient-domain machine learning [164]

Lastly, a series of methods exist that utilize convolutional neural networks such as the continuous-filter convolutional neural network of SchNet [165], and the method we will focus on in detail in the next section, Covariant Convolutional Neural Networks [166]

Much attention has been paid to molecular machine learning frameworks and therefore some architectures are not able to extend to periodic systems. When moving to crystals a series of complications arrive due to periodic images, ambiguity in conventional cell choice, and extensively.

5.2 Covariant Convolutional Neural Networks

While we found tremendous success with the use of BPNN with our system, the featurization is not perfect. The process of hyperparameter selection is tedious and required either guess and check or at best many individual models to be trained and tested with the assistance of Bayesian optimization. This fixed set of fingerprints may work for the current data set but they may need to be re-selected and optimized if the data set grows and changes. Particularly if the characteristic interactions in data drastically change. Additionally, the atomic features are completely invariant to rotations, therefore, losing information of directionality in the material. If the prediction of the BPNN was to be extended to vector quantities such as spontaneous polarization of tensor quantities like those related to nuclear magnetic resonance, this rotational information would need to be considered. We present here a promising new approach to atomic featurization that enables not only the automatic learning of atomic features through convolutions, but also is rotationally covariant and therefore is intelligent to rather than ignorant of vector and tensor valued physics.

Toward this goal, a class of symmetric equivariant (covariant) neural network frameworks have been developed. In these methods, the features transform in a mathematically well-behaved manner such that the symmetry transformation can be applied before the featurization of the material or to the featurization itself and yield the same results. This is to say if $f(x)$ is a function that maps an input vector x of coordinates and chemical identities of the atoms in a material to feature space, and T_g is some transformation of the material, such as a translation or rotation that will yield the same energy of the material, then the featurization is covariant if

$$f(T_g * x) = T_g * f(x) \tag{5.1}$$

Convolutional Neural Networks (CNNs) for the use of image recognition perfectly demonstrates this concept. Within CNN models, an input image is transformed to the brightness of red, green, and blue (RGB) in each of these channels, learnable stencils are learned that convolve a collection of pixels in order to pick out patterns in the image and learn the "pixel environment" of an image. These stencils themselves, are both translationally equivariant, and given that they are learned, this architecture automatically learns the features in a systematic way.

Unfortunately, conventional CNNs are not covariant to rotations, which is an important transformation within physics. To extend this concept, Spherical CNNs were developed by T. Cohen and R. Kondor [167, 168]. By transforming to a generalized Fourier Space in which the features and transformation are decomposed in terms of the irreducible representation of the group of transformations to which the architecture is covariant, the model can more computationally efficiently evaluated [167]. Additionally, the permutation invariance of the atom labelings has been well studied in the context of learning material properties in message-passing graph networks [169].

Ideally, an architecture would unify translation, rotation, and permutation invariance in an architecture that requires little hand-picking of features. The most promising application of the unification of these concepts to the featurization of materials, the *COvaRiant MOleculaR Artificial Neural neTwork (Cormorant)* [166]. This architecture uses the spherical tensor objects of Clebsh Gordon nets. The method of connecting and manipulating these tensor objects marries conventional CNN and message passing graph neural network concepts with two distinct activation types. The first is vertex activations, F_i , that resemble conventional convolutional neural networks where each vertex in the CNN first represents a single atom and moving up in the architecture, the vertex represents a collection of atom in a neighborhood, until the last level a single vertex represents all atoms in the material. The second kind of activation is an edge type, G_{ij} . This represents a weighted graph connecting all vertices of the CNN. The weights of this graph and therefore the value of G_{ij} , depends on on the values of the vertex activations and with the inclusion of ration invariance that is absent from other graph neural networks can encode the relative positions of atoms and angles of interactions as well as distances and chemical identities. The rotational invariance is maintained by expanding the activations to spherical tensors as $F_i = (F_i^0, \dots, F_i^{\ell_{\max}})$ and $G_{ij} = (G_{ij}^0, \dots, G_{ij}^{\ell_{\max}})$, where $F_i^\ell \in \mathbb{C}^{(2\ell+1) \times n_\ell}$ is a spherical tensor of order ℓ , and n_ℓ is the multiplicity (number of channels) of the tensor.

As with Clebsh Gordon nets, the covariance is maintained with a carefully chosen non-linearity. that is dictated by the structure and algebraic concepts of the ration group and thus the Clebsch-Gordan product (CG product), \otimes_{cg} , of two tensors is chosen and defined as

$$[A_{\ell_1} \otimes_{\text{cg}} B_{\ell_2}]_\ell = \bigoplus_{\ell=|\ell_1-\ell_2|}^{\ell_1+\ell_2} C_{\ell_1 \ell_2 \ell} (A_{\ell_1} \otimes B_{\ell_2}) \quad (5.2)$$

where \otimes denotes a Kronecker product, and $C_{\ell_1 \ell_2 \ell}$ are the well known Clebsch-Gordan coefficients [167].

The activations F_i^s at level s are chosen to be

$$F_i^s = \left[F_i^{s-1} \oplus (F_i^{s-1} \otimes_{\text{cg}} F_i^{s-1}) \oplus \left(\sum_j G_{ij}^s \otimes_{\text{cg}} F_j^{s-1} \right) \right] \cdot W_{s,\ell}^{\text{vertex}} \quad (5.3)$$

$$F_i^s = (F_i^{s-1} \otimes_{\text{cg}} F_i^{s-1}) \cdot W_{s,\ell}^{\text{vertex}} \quad (5.4)$$

where \oplus denotes concatenation and $W_{s,\ell}^{\text{vertex}}$ is a linear mixing layer that acts on the multiplicity index.

The edge activations are chosen to have the form of $G_{i,j}^{s,\ell} = g_{ij}^{s,\ell} \times Y^\ell(\hat{\mathbf{r}}_{ij})$, where \mathbf{r}_{ij} is the relative position vector pointing from atom i to atom j . The scalar-valued edge terms are then given by

$$g_{ij}^{s,\ell} = \mu^s(r_{ij}) \left[\left(g_{ij}^{s-1,\ell} \oplus (F_i^{s-1} \cdot F_j^{s-1}) \oplus \eta^{s,\ell}(r_{ij}) \right) \cdot W_{s,\ell}^{\text{edge}} \right] \quad (5.5)$$

with $\mu^s(r_{ij})$ a learnable mask function, $\eta^{s,\ell}(r_{ij})$ a learnable set of radial basis functions, and $W_{s,\ell}^{\text{edge}}$ a linear layer along the multiplicity index.

The architecture is iterated for $s = 0, \dots, s_{\text{max}}$. Finally, at the last layer of *Cormorant*, the $\ell = 0$ component of the output, which represents a rotationally invariant quantity is used to predict the energy of the material.

We have extended this work which has shown tremendous success in predicting energy, polarizability, and other quantities for molecules [166] to periodic systems for the study of crystals including the battery materials here. We have additionally, add the capability to predict forces as from the negative of the analytical gradient of the predicted energy with respect to the position of that atom. For molecules, the algorithm samples over every neighbor j , of atom i . In periodic systems, however, there exist infinitely many copies of atoms j . To address this, an interaction cutoff distance is used and the input data is automatically augmented such that only a single image of atom j appears within the cutoff distance of atom i . We then use the minimum image convention to calculate the smallest distance between i and j .

Additionally, when moving the periodic systems, the extensiveness of the architecture must be properly considered. The molecular architecture of *Cormorant* took each atom feature vector as an input to a final multi-layer perceptron (MLP) to learn the mapping from the fingerprints to the property to be predicted. As the MLP is a nonlinear operation, this architecture with not be extensive by default. For the crystal system, both an atom-wise MLP on the atom features as seen in the BPNN, and linear regression with no bias term can be used and are implemented.

To test the ability of architecture to featurize materials, a subset of the data used in Chapter 4 was used to train an architecture with 12 channels and various number Clebsch Gordon layers testing both with an output MLP and without and output MLP. The results of these tests are shown in Table 5.1. We see that surprisingly, the performance of the overall model is improved when linear regression is performed on the learning representation of the materials from the CG layers. This suggests that the automatic pattern recognition of the convolutional neural network is learning physically significant features on which the energy directly depends.

For the convenience of collecting data and easy transference of the final calculator, the periodic *Cormorant* implementation was interfaced with the Atomic Simulation Environment package. This package not only carries the ASE-database framework for easy data collection and manipulation but provides a python based calculator for the *Cormorant* model. As ASE is already interfaced with a large collection of density functional theory, molecular dynamics, and machine learning codes, this improves the ease of use and reproducibility of the code.

This implementation is just the beginning of the possibilities of this architecture. The rotationally covariant nature of *Cormorant* could allow for the training of the stress tensor and therefore ultra-accurate cell shape optimization. It could also have applications to ferroelectrics by learning the spontaneous polarization of materials.

	Layers	Training (eV)		Test (eV)	
		MAE	RMSE	MAE	RMSE
output MLP	5	0.0028	0.0045	0.0027	0.0045
	6	0.0026	0.0034	0.0024	0.0035
	7	0.0048	0.0063	0.0045	0.0067
no output MLP	5	0.0018	0.0031	0.0018	0.0027
	6	0.0019	0.0033	0.0019	0.0029
	7	0.0013	0.0017	0.0014	0.0021

Table 5.1: A comparison of training results on a subset of the NMC data from Chapter 4. The use of an output multi-layer perceptron (MLP) was tested versus linear regression on the learned feature of the convolutional layers.

Chapter 6

Singlet Oxygen Evolution In Battery Cathodes

The inadvertent generation of the highly reactive excited spin state of molecular oxygen known as singlet oxygen has emerged as a clear adversary in a range of battery technologies leading to parasitic current and electrolyte degradation. Having been observed in Li-air, Na-air, and NMC cathodes, an understanding of the mechanism of generation could inform electrolyte design on a broad front of battery science. We present here an analysis of singlet oxygen production from both free super-oxide anions in solution as well as alkali-superoxide ion pairs through charge disproportionation. Using Marcus theory of kinetics we understand both a mechanism for the direct production of the singlet excited state as well as the effect of the electrolyte properties on the reaction rate for both triplet and singlet oxygen production. Through this, we can rationalize the experimentally seen onset voltage for singlet oxygen production in each of these chemistries.

The key insight we present is that the cation associated disproportionation reaction commonly investigated in literature will favor triplet ground state triplet oxygen production. The free superoxide reaction on the other hand will favor singlet production due to the much more negative reaction energy. It is therefore the dissociation reaction and interplay between the cation associated and free superoxide that will determine singlet oxygen yield. We also provide insight into the decomposition of lithium carbonate (Li_2CO_3), a well known residual compound from cathode synthesis. Through DFT calculations, a mechanism similar to that seen in lithium and sodium air batteries is understood therefore unifying the understanding of the mechanism for a common degradation scheme.

6.1 Excited States of Molecular Oxygen

The electron structures of the various state of molecular O_2 are determined by the two electrons in the π^* anti-bonding molecular orbital. The addition of the quantum

spin of two electrons creates two energy levels, the singly degenerate spin state with a total spin quantum number of 0 and the triply degenerate spin state with a total spin quantum number of 1. In the case of the oxygen molecule, the relativistic interaction of the spin with the molecular axis angular momentum (called the spin-axis interaction) is large compared to the difference in rotational energy levels of the molecule and therefore obeys Hund’s rule [170]. Hund’s rule suggests that the electric configuration with the greatest total spin will be the ground state, correctly predicting the triplet spin state to be the true ground state of oxygen. Oxygen is therefore unique compared to other homonuclear diatomic molecules that exist in a spin-singlet ground state.

The two lowest energy levels of oxygen are thus, ${}^3\Sigma_g^-$, which we refer to as triplet oxygen, and $({}^1\Delta_g)$, which we refer to as singlet oxygen and is 0.98 eV higher in energy [171]. Finally, a second spin-singlet excited state, $({}^1\Sigma_g^+)$, exists 1.64 eV about the ground state [171]. Due to the spin transition rules, forbidding a $\Delta S = 0$, the transition from singlet to the triplet ground state will rarely occur directly. This leads to the extremely reactive nature of singlet oxygen, especially with organic molecules such as the non-aqueous electrolytes used in all of the batteries considered in this chapter.

The mechanistic understanding within this paper leverages insight from singlet oxygen production in biology, which has been seen in water in the presence of superoxide (O_2^-) and protons in dimethylsulfoxide (DMSO) [172], and in acetonitrile (MeCN) with superoxide and ferrocenium cations [173], both of which happen to be common battery electrolytes. In both of these cases, the disproportionation (also termed dismutation) of superoxide was proposed as the method of singlet production: $\text{O}_2^- + \text{O}_2^- \longrightarrow \text{O}_2 + \text{O}_2^{2-}$, and later this mechanism was confirmed [174, 175]. We, therefore, wish to understand the kinetics of this reaction both with free superoxide and superoxide in the presence of Li^+ in various electrolytes.

6.2 Marcus Theory of Electron Transfer

In the reaction of interest, we assume an outer shell electron transfer occurs between a donor species and an acceptor species in which no bonds are created or broken. The rate constant of this reaction, which has some energy barrier, ΔG_D^* , is expressed by the Arrhenius relation:

$$k = Z \exp\left(-\frac{\Delta G_D^*}{kT}\right) \quad (6.1)$$

where the pre-exponential factor Z is the theoretical number of collisions between two neutral species in solution and has been previously tabulated to be on the order of $10^{11} \text{M}^{-1} \text{s}^{-1}$ [176]. The exact value of this pre-exponential factor is not critical as we will be comparing rate constants of the same reactants. The reaction barrier, ΔG_D^* , is determined by the intersection of the potential energy surface (PES) of the

product and reactants. In the theory of Marcus, these potential energies are assumed to be harmonic, leading to an expression for the reaction barrier in terms of the standard free energy change, ΔG^0 , and the energy to be dissipated for an electron on the acceptor PES at the equilibrium position of the donor PES, to the equilibrium position of the acceptor PES. [176] This energy dissipation term is known as the reorganization energy λ .

$$\Delta G_D^* = \frac{1}{4\lambda} (\Delta G^0 + \lambda)^2 \quad (6.2)$$

This reorganization energy consists of two terms. The first is the reorganization of the nuclei between the reactant and the products state and depends on the vibrational properties of the molecules involved. For simplicity, we neglect this term for this analysis and assume the nuclear positions are constant for both the initial and final states. The second component of the reorganization energy is the electrostatic rearrangement of the electrolyte molecules in response to the change in charge of the molecules being solvated. The electrolyte can be treated as a dielectric medium with static relative permittivity (dielectric constant), D_s , and optical relative permittivity, D_{op} , equal to the index of refraction squared. This dielectric medium is solvating two spheres of radius a_1 and a_2 separated by R to give a reorganization energy upon the transfer of Δe charge of [176]

$$\lambda = \frac{(\Delta e)^2}{4\pi\epsilon_0} \left(\frac{1}{D_{op}} - \frac{1}{D_s} \right) \left(\frac{1}{2a_1} + \frac{1}{2a_2} - \frac{1}{r} \right) \quad (6.3)$$

Within this work, we assume the two reacting species have a radius of $a = 2 \text{ \AA}$, separated by a distance of $R = 3.5 \text{ \AA}$ based on the distances seen in first-principle calculations of LiO_2 disproportionation reaction in the gas phase [178].

One of the most remarkable predictions of Marcus theory and Eq. 6.2 is the prediction of an absolute minimum in the reaction barrier as a function of the standard free energy change. A simple maximization of this equation shows that this maximum occurs at $\Delta G^0 = -\lambda$. That means that as the reaction becomes increasingly exergonic and the thermodynamic driving force increases, the rate of the reaction will not continue to increase but will instead reach a peak and then begin to decrease in what is known as the inverted region. More interestingly, this means that as the standard free energy of reaction to the ground state becomes more negative, the reaction rate of an excited product state could be larger than that of the ground state product reaction rate. To further illustrate this point, the reaction rate constant, k , is plotted in Fig. 6.1 for both the production triplet and the first singlet excited state in the disproportionation reaction of superoxide.

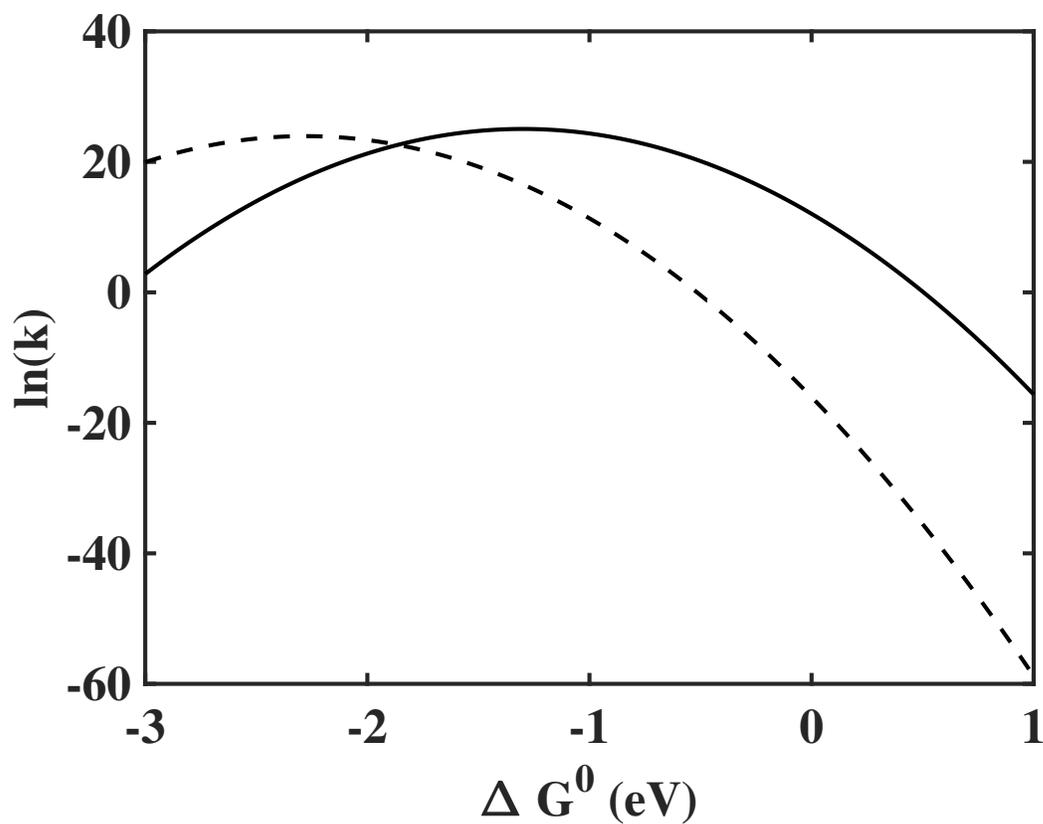


Figure 6.1: The log of the predicted rate of oxygen production in the singlet (dotted line) and triplet (solid line) spin state as a function of the change in Gibbs energy for the triplet ground state reaction with a reorganization energy of $\lambda = 1.3\text{eV}$. Note the inverted region in both curves, demonstrating that the rate does not continue to increase as the thermodynamic driving force increases. Thus the rate of singlet ultimately overtakes the rate of triplet oxygen production at a high driving force. [Reprinted with permission from [177]. Copyright 2020 American Chemical Society]

6.3 Solvent Effects on Reaction Energy

We focus on lithium and will later qualitatively extrapolate the result to predictions for related to reactions involving sodium. For a thorough treatment of the problem, we consider both the disproportionation reaction of free O_2^- and LiO_2 ion pair dissolved in solution. The two reaction energies of interest are

$$\Delta G^0 = G_{O_2^{2-}}^{(sol)} + G_{O_2}^{(sol)} - 2G_{O_2^-}^{(sol)} \quad (6.4)$$

$$\Delta G^0 = G_{Li_2O_2}^{(sol)} + G_{O_2}^{(sol)} - 2G_{LiO_2}^{(sol)} \quad (6.5)$$

For the first equation, we can use the gas phase molecular oxygen electron affinity of 0.44 eV [179] and the previously estimated gas-phase reaction enthalpy for the reaction $O_2 + 2e^- \longrightarrow O_2^{2-}$ of 4.77 eV [180]. Since within the Born approximation of solvation, the solvation energy is proportional to the charge of the species squared, we assume the solvation energy of O_2^{2-} to be four times that of O_2^- . For the second equation, we use the previously computed gas phase ΔG^0 for the full reaction of 0.73 eV [178]. We also estimate the solvation energy of Li_2O_2 to be approximately 1.37 times that of LiO_2 as is supported by previous calculations of the solvation energies of these species [181] and can be understood by the increase in charge localization of Li_2O_2 increasing the solvation energy. To understand the effect of the solvent properties on these reactions, we approximate the solvation energies of both the product and reactant species as functions of the Gutmann donor and acceptor number of the electrolyte. It has been previously shown that the solvation energies of dissolved Li^+ and O_2^- are linear functions of the donor and acceptor number respectively [182] and we, therefore, take these fits. We also approximate the solvation energy of the neutral LiO_2 and Li_2O_2 species as a function of the sum of donor and acceptor number as seen in Fig. 6.2(a-c).

Using these approximations, we evaluated Equations 6.4 and 6.5 as well as the ultimate reaction rate constant for a series of electrolytes in Table 6.1. We find that the free superoxide and cation associated superoxide reactions predict different rate constants that support the kinetic favorabilities of the different spin state of the oxygen produced. While the $Li^+ - O_2^-$ ion pair, yields a high rate constant for triplet, the free O_2^- yields the highest rate constant for singlet.

To predict the ultimate reaction rate, ν , for each reaction, the concentration of reactants is required. The concentration of each of the reactants is determined by the equilibrium constant of the dissociation reaction of the superoxide with lithium $LiO_2 \longrightarrow O_2^- + Li^+$

$$K^{eq} = \frac{[O_2^-][Li^+]}{[LiO_2]} = \exp \frac{-\Delta G_{dis}}{kT} \quad (6.6)$$

The change in free energy of this reaction has been calculated previously for four electrolytes and fit to a linear function of the combined solvation energy of

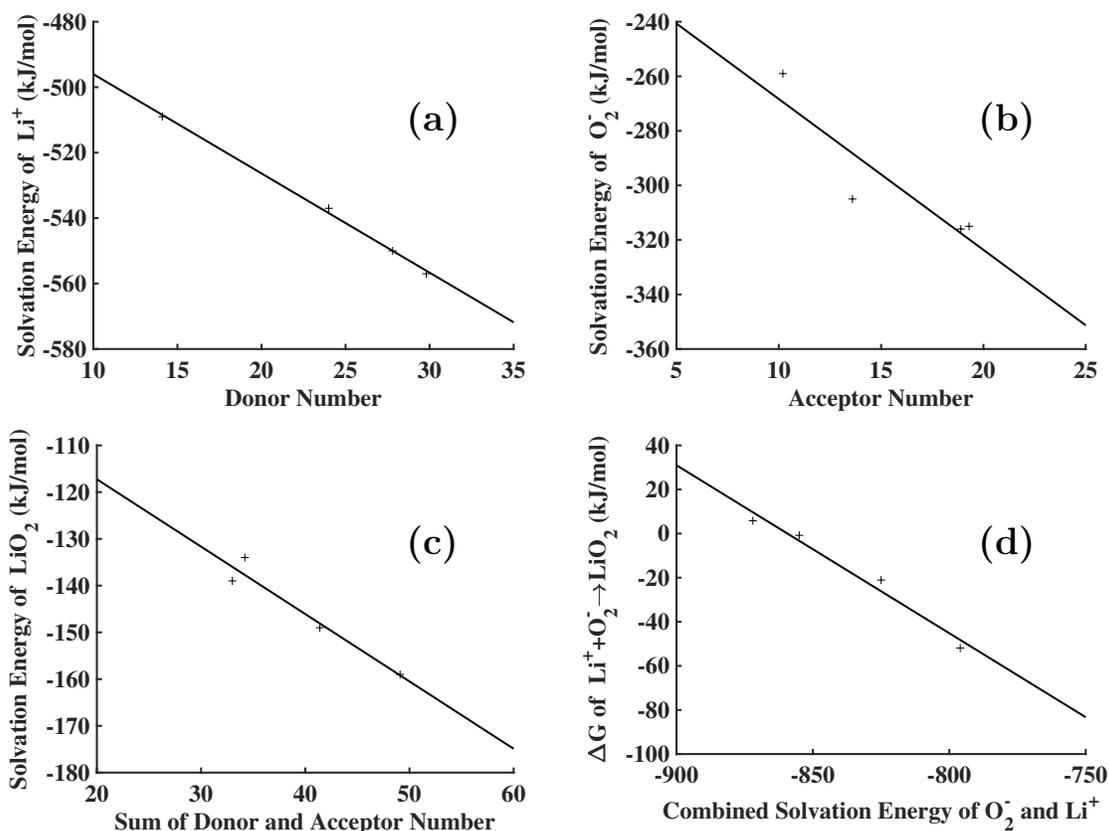


Figure 6.2: The linear fits for (a) the solvation energy of Li^+ as a function of donor number of the solvent, (b) the solvation energy of O_2^- as a function of acceptor number of the solvent, (c) the solvation energy of LiO_2 as a function of the sum of donor and acceptor number of the solvent, and (d) The dissociation energy of LiO_2 into Li^+ and O_2^- as a function of the combined solvation energies of the free ions. All numbers are for dimethoxyethane (DME), dimethylsulfoxide (DMSO), acetonitrile (MeCN), and dimethylacetamide (DMA) taken from Kwabi *et. al.* in Reference [182].

O_2^- and Li^+ as shown in Fig. 6.2(d). The concentration the lithium is controlled experimentally by the concentration of the lithium salt within the electrolyte, and is set here to $[Li^+] = 0.1M$. Now, from the reaction rates to produce oxygen in a spin state s , $\nu_s^{LiO_2} = k[LiO_2]^2$ and $\nu_s^{O_2^-} = k[O_2^-]^2$, and Eq. 6.6, we compute the predicted fraction of singlet oxygen produce as

$$f_{1O_2} = \frac{\nu_1^{LiO_2} + \nu_1^{O_2^-}}{\nu_1^{LiO_2} + \nu_3^{LiO_2} + \nu_1^{O_2^-} + \nu_3^{O_2^-}} \quad (6.7)$$

Shown as two limiting cases for the range of reorganization energies seen in battery electrolytes is the predicted f_{1O_2} for dimethoxyethane (DME) with low reorganization energy and acetonitrile (MeCN) with high reorganization energy in Fig. 6.3 From this figure, we see that solvents with high donor and high acceptor are predicted to release a higher proportion of singlet oxygen. There are three reasons compounding reasons for this prediction. Firstly, a solvent that is in general highly solvating with both a high acceptor and high donor number will favor the dissociation of LiO_2 into free ions in solution. This increase in O_2^- will then drive the singlet favoring superoxide disproportionation reaction. Next, the solvation energy of O_2^{2-} will increase with an increasing acceptor number of the electrolyte solvent. Thus the energy from equation 6.4 will be more negative driving the kinetics farther into the region favoring singlet production. If the acceptor number is too low, however, the free superoxide reaction will not occur independently of both the concentration of superoxide and the donor. This result proposes that the acceptor number is more important than the donor number which has been used as the sole descriptor in the studying superoxide disproportionation, as even for low donor numbers, we predict an acceptor number at which singlet oxygen will still be produced. And finally, a highly solvating electrolyte will slow the triplet favoring cation associated reaction. Looking at eq. 6.5, the solvation energy of two LiO_2 is higher than that of the single Li_2O_2 produced and thus an increase in solvation will stabilize the reactants and slow this reaction.

It should be noted that due to the simplicity of the Marcus theory of kinetics and the assumptions for solvation energies a simple linear functions of electrolyte donor and acceptor number, these predictions are more qualitative. Additionally, these predictions can only comment on the relative production of singlet oxygen versus triplet. The plots in Fig. 6.3 do not comment on the absolute production of oxygen. In fact in high donor and acceptor number solvent, such as DMSO, the total amount of oxygen release could be low. This analysis also ignores the effect of singlet quenching. As superoxide is a well-known quencher of singlet oxygen [183], at high concentrations, it would both produce and mediate the harmful effects of singlet oxygen and therefore the amount seen would vary greatly from the theoretical predictions here. Despite the approximations of this analysis, we rationalize many experimental findings related to the effect of cations on superoxide disproportionation and singlet oxygen production as well as explain the onset potentials for singlet oxygen release in batteries.

Table 6.1: Predicted ΔG^0 to form ground state O_2 for the free superoxide and cation associated superoxide reactions from Equations 6.4 and 6.5 respectively and the corresponding rate constant in various solvents: tetraethylene glycol dimethyl ether (TEGDME), dimethoxyethane (DME), dimethyl sulfoxide (DMSO), dimethylacetamide (DMA), dimethylformamide (DMF), and acetonitrile (MeCN). The solvent reorganization energy was calculated using Eq. 6.3 using $a=2.0$ Å and $R=3.5$ Å for the radius of each reactant and separation respectively. Data reproduced in part from Ref [177].

Solvent	Reactant	ΔG^0 (eV)	Spin State	Rate Constant
TEGDME	LiO ₂	0.10	$^3\Sigma_g^-$ $^1\Delta_g$	2.1×10^5 1.7×10^{-8}
	O ₂ ⁻	-1.74	$^3\Sigma_g^-$ $^1\Delta_g$	2.1×10^9 9.1×10^9
DME	LiO ₂	0.16	$^3\Sigma_g^-$ $^1\Delta_g$	2.2×10^4 1.2×10^{-9}
	O ₂ ⁻	-1.71	$^3\Sigma_g^-$ $^1\Delta_g$	8.5×10^9 4.5×10^9
DMSO	LiO ₂	0.30	$^3\Sigma_g^-$ $^1\Delta_g$	2.0×10^2 5.0×10^{-12}
	O ₂ ⁻	-2.75	$^3\Sigma_g^-$ $^1\Delta_g$	4.1×10^4 6.3×10^9
DMA	LiO ₂	0.23	$^3\Sigma_g^-$ $^1\Delta_g$	5.9×10^2 6.6×10^{-11}
	O ₂ ⁻	-2.10	$^3\Sigma_g^-$ $^1\Delta_g$	2.8×10^9 1.4×10^{10}
MeCN	LiO ₂	0.15	$^3\Sigma_g^-$ $^1\Delta_g$	4.2×10^2 4.3×10^{-10}
	O ₂ ⁻	-2.71	$^3\Sigma_g^-$ $^1\Delta_g$	6.8×10^7 2.3×10^{10}

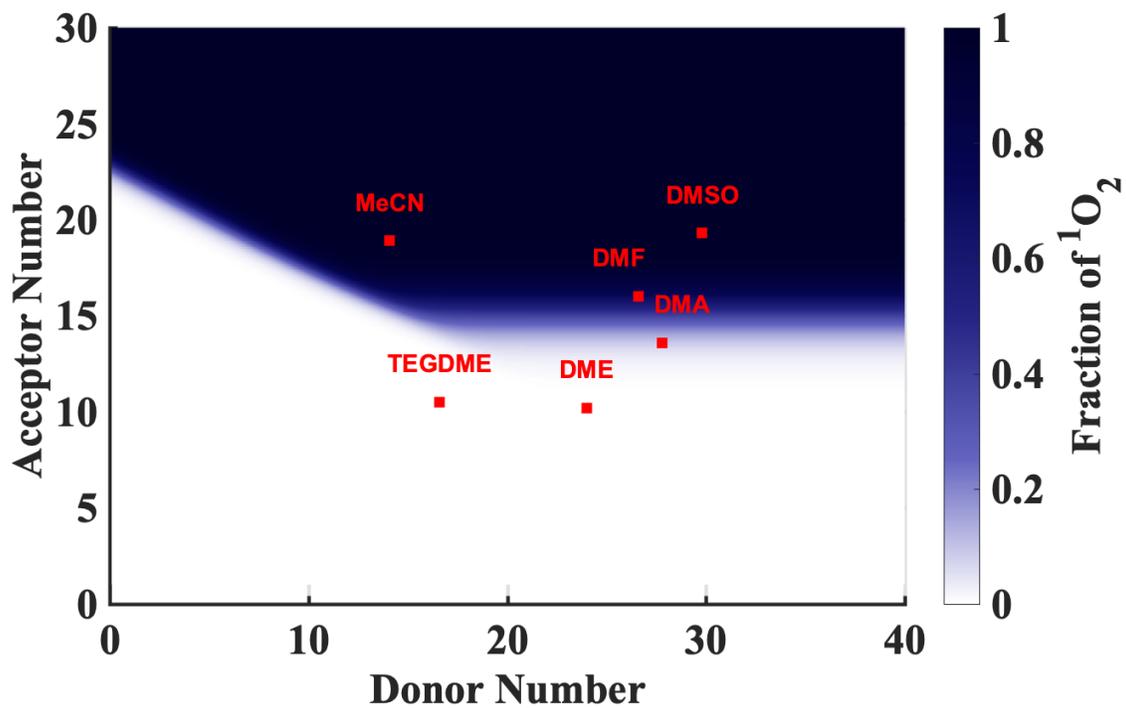
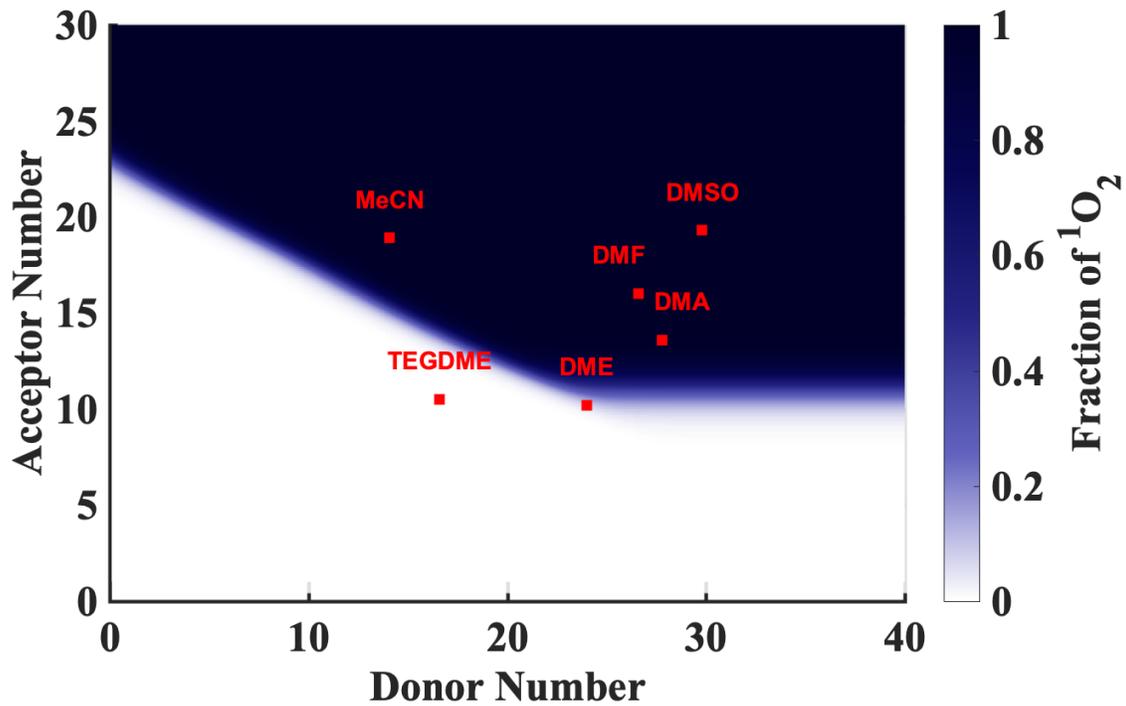


Figure 6.3: both ($^1\Delta_g$) and the ($^1\Sigma_g^+$) spin singlet states considered. (a) $\lambda = 1.19$ eV to represent DME and (b) $\lambda = 1.62$ eV to represent MeCN. [Reprinted with permission from [177]. Copyright 2020 American Chemical Society]

6.4 Understanding of Experimental Observation

Experimental investigation of the cation driven disproportionation reaction of superoxide compared the rate of singlet production in the presence of Li^+ and Na^+ [184]. For Na^+ , they found an increase in the proportion of singlet oxygen with a decrease in overall oxygen release compared to Li^+ . This is immediately explained within our analysis as there would be more dissociation of NaO_2 since Na^+ is a weaker Lewis acid than Li^+ . That is, it has less of a tendency to accept electrons and will form a weaker bond with the O_2^- . The fraction of singlet oxygen would, therefore, increase due to the increase in O_2^- in solution. The dominant reaction, however, would continue to be the cation associated superoxide reaction. The rate of this for Na however would be lower than the Li equivalent since Na_2O_2 is less stable than Li_2O_2 . Another experimental work on the disproportionation of NaO_2 which did not attempt to detect singlet yields noticed that triplet oxygen was only released for low donating solvents [185]. As was previously discussed, the high solvating properties of an electrolyte could stabilize NaO_2 and prevent disproportionation, therefore leading to less triplet oxygen released. While these two experiments on isolated superoxide systems can give insight, we look to understand the realization of the mechanism within battery operation.

6.4.1 Lithium Air Batteries

Lithium-air batteries are a promising technology due to their extremely high specific energy. This technology removes the weight of all intercalation electrodes, using pure Li-metal as the anode, and the O_2 from the air as the cathode. In the discharged state, the weight of the battery consists only of the lithium metal, the electrolyte, and the current collectors. The operation of the battery works on the following reaction.



While an exciting concept, in practice, issues related to reversible electrodeposition of the Li-metal anode as well as passivation of the cathode present challenge to long term cycling. Additionally, parasite current from side reaction and electrolyte degradation lead to poor cycling efficiency and cell death. Much of this side-current and electrolyte decomposition has been linked to the production of singlet oxygen [186, 187, 188] In fact, singlet has been observed during charge, discharge, and rest of lithium-air ($\text{Li}-\text{O}_2$) batteries. [186]

During charge, the largest rate of singlet is observed with an onset potential of 3.5V [186]. This voltage was previously believe to be due to the direction decomposition during charge $\text{Li}_2\text{O}_2 \longrightarrow 2\text{Li}^+ + \text{O}_2 + 2\text{e}^-$. The reversible electrochemical potential of this reaction to produce ground state O_2 is 2.96 V, while as discussed before, the singlet oxygen state is 0.98 eV above triplet. While this would imply, the thermodynamic onset to be 3.45V, close to that seen experimentally, there is no

mechanistic reason for the kinetic preference of singlet over the much more favorable triplet state and a direct release is improbable. We propose that the onset is due to the dissolution of LiO_2 from the surface. $\text{LiO}_2^* \longrightarrow \text{Li}^+(\text{sol}) + \text{O}_2^-(\text{sol})$. Thus the solvent-mediated disproportionation reaction we propose will then occur. Additionally, the onset of when the LiO_2 is controlled by the potential dependent termination of the surface. Previous computational work has shown that at 3.4 V, the most stable surface becomes the O-rich3 (1-100) termination [189]. This means that as the potential raises, the surface is stripped in a way that exposes this new oxygen-rich surface. The LiO_2 can then more easily dissolve in solution and participate in the solution mediate reactions discussed previously.

In parallel, during discharge, less singlet oxygen is seen and even less during rest [187]. The mechanism for discharge proceeds through two single-electron steps to react the Lithiums one by one to adsorbed species on the surface. $\text{Li}^+ + \text{e}^- + \text{O}_2^* \longrightarrow \text{LiO}_2^*$ followed by $\text{Li}^+ + \text{e}^- + \text{LiO}_2^* \longrightarrow \text{Li}_2\text{O}_2$. For the second step, it is well known that a solution mediate reaction competes with the surface-mediated reaction. This reaction is the disproportionation of two Li_2O_2 $2\text{LiO}_2 \longrightarrow \text{Li}_2\text{O}_2 + \text{O}_2$ Thus this competing solution reaction, more specifically the free O_2^- analog, is the reaction from which singlet will be released. The competition with the surface-mediated reaction then accounts for the lower amount of singlet seen.

6.4.2 Sodium Air Batteries

Another promising battery chemistry is that of sodium air. The relative abundance of Na versus Li as well as the promise to form more stable superoxide makes this a promising alternative to lithium-air despite its slight decrease in energy density. Just as with lithium-air, however, singlet oxygen production and it's corresponding electrolyte degradation has been seen during charge, discharge and rest [190]. With a reversible potential of 2.27V, the charge reaction proceeds through $\text{NaO}_2 \longrightarrow \text{Na}^+ + \text{O}_2 + \text{e}^-$. Again from a purely thermodynamic analysis, the singlet excited state would have a thermodynamic onset of 3.25V. Again, however, there is no kinetic reason for the preference of singlet during release and therefore this mechanism is again improbable. Furthermore, singlet is seen from the start of charging and at all charging potentials as low as 2.27V, suggesting the release of singlet is not direct. We again propose a change in surface termination leading to the dissolution of O_2^- into solution and the proposed solution mediated process. During discharge the stable surface is the stoichiometric 100 termination forming perfectly cubic crystal. [191] At all reasonable charging potential the corners of the surface are decomposed first due to a lower kinetic barrier of oxidation forming the more oxygen-rich 111 and 110 surface [191]. Just as with the lithium-air system, this leads to the dissolution of NaO_2 and followed by disproportionation. During discharge, just as with lithium, the rate of singlet lower. And the discharge process $\text{Na}^+ + \text{e}^- + \text{O}_2^* \longrightarrow \text{NaO}_2^*$ will again compete with solution mediated disproportionation. This leads to Na_2O_2 on

the surface that is experimentally seen but was previously poorly understood from the poor thermodynamics of Na_2O_2 [192, 190].

6.4.3 Nickel Rich Lithium-Ion Batteries

The clear analogies between the Li-air and Na-air batteries lead to the easily understandable unification of singlet production within these two systems. The reactions and chemistry of Li-ion batteries with intercalation rather than an O_2 cathode may appear distinct at first glance. However, singlet oxygen has also been observed in Ni-rich NMC batteries [4] and in general, electrolyte degradation has been attributed to oxygen release [193]. This degradation of the organic would likely lead to CO_2 release. The CO_2 release seen during operation, however, has complex origins. It was experimentally seen that a majority of the CO_2 released during the first cycle of Ni-rich cathodes can be linked using isotopic tracing to residual Li_2CO_3 from the synthesis of the NMC cathode [194]. Due to the tendency of Ni to occupy the Li site, to properly synthesize Ni-rich cathodes, excess lithium is required [195] and this excess of lithium leads to residual Li_2CO_3 . The decomposition of Li_2CO_3 is proposed to occur through $2\text{Li}_2\text{CO}_3 \longrightarrow 4\text{Li}^+ + 4\text{e}^- + 2\text{CO}_2 + \text{O}_2$ with a theoretical reversible potential of 3.82 V that is confirmed experimentally. During this decomposition, however, there is a mysterious lack of oxygen detected. Later it was experimentally measured, that the decomposition of Li_2CO_3 releases predominately singlet oxygen explaining the lack of detection of molecular oxygen. Therefore, a portion of the singlet oxygen in NMC could be explained by the decomposition of Li_2CO_3 . The direct release of singlet can be ruled out similarly as before as thermodynamically the singlet would release at 4.05 V. During the experimental detection during lithium carbonate decomposition, the release however begins at 3.82 V as expected for ground state O_2 . To probe the decomposition mechanism, we performed DFT calculations to understand the steps of decomposition. Starting from the most stable 001 surface of Li_2CO_3 seen in 6.4(a-b) we performed trial removals of O_2 , or CO_2 . If the chemical removal of either of these was not energetically favorable, instead a Li^+ and corresponding e^- were removed to simulate the electrochemical oxidation under an applied potential U . The U required to perform this removal is such that the positive change in free energy of the removal reaction is balance by the electrostatic interaction of the released electron with the potential U .

$$eU = -\Delta G \tag{6.9}$$

The results of the decomposition, shown in Fig. 6.4(c) find that after the removal of five Li^+ , two CO_2 spontaneously release from the surface. What was left on the surface, was a $\text{LiO}_2\cdot$ like moiety. Using a Bader charge analysis [196], the oxidation states of the oxygen atoms were confirmed to be each $-1/2$ [177] suggesting that together they form a O_2^- like moiety. This LiO_2 can then dissolve from the surface and participate in the disproportionation reaction previously discussed. This reaction

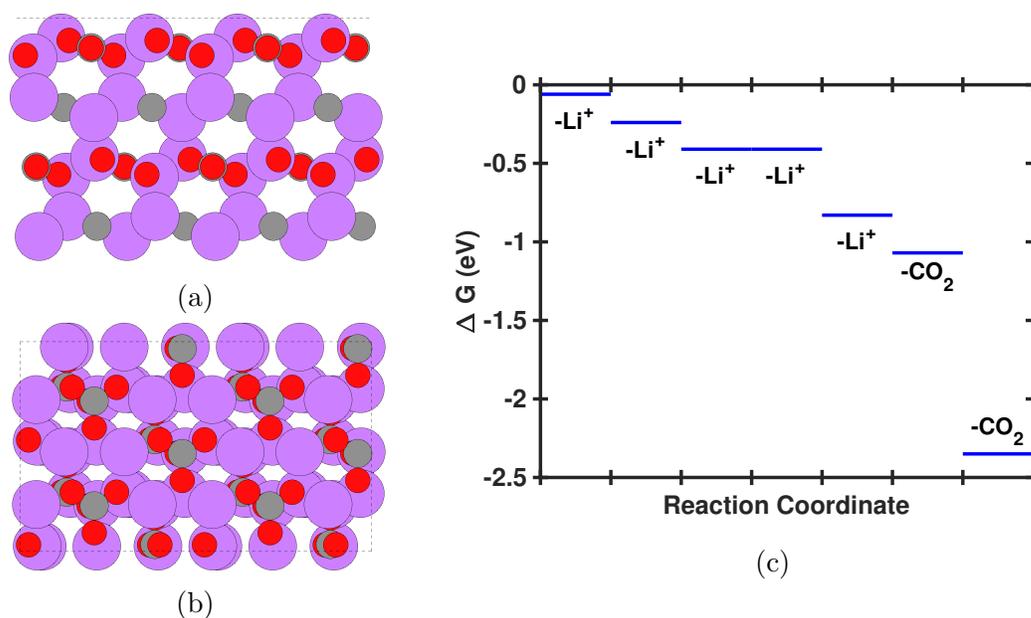


Figure 6.4: Calculated reaction pathway (c) for the decomposition of Li_2CO_3 surface shown from the side (a) and from above (b). Lithium atoms are shown in purple, oxygen in red, and carbon in gray. [Reprinted with permission from [177]. Copyright 2020 American Chemical Society]

will produce Li_2O_2 that will immediately precipitate to the surface. As the redox potential of Li_2O_2 is well below the potential at which the Li_2CO_3 is decomposed, it will immediately oxidize, leaving behind another adsorbed $\text{LiO}_2\cdot$ that can also eventually participate in disproportionation. A schematic for this full decomposition and singlet oxygen production mechanism in lithium carbonate as well as in Li-air seen in Fig. 6.5

This mechanism explains the singlet oxygen release around 3.8V in NMC cathodes. There is, however, the release of singlet oxygen seen at higher potentials of around 4.5V. The release was seen in a variety of NMC chemistries all with different onset potentials. The onset did however always occur at near 80% state of charge suggesting a common instability due to destabilization from the missing Li. Additionally, this onset can be induced in a purely thermal way. Cathodes were charged to 80% state of charge and then removed from the cell. They were then placed in an argon environment and heated leading to a detectable release of singlet oxygen. [4] Therefore the solution mediated mechanism cannot explain this direct release of lattice oxygen as singlet.

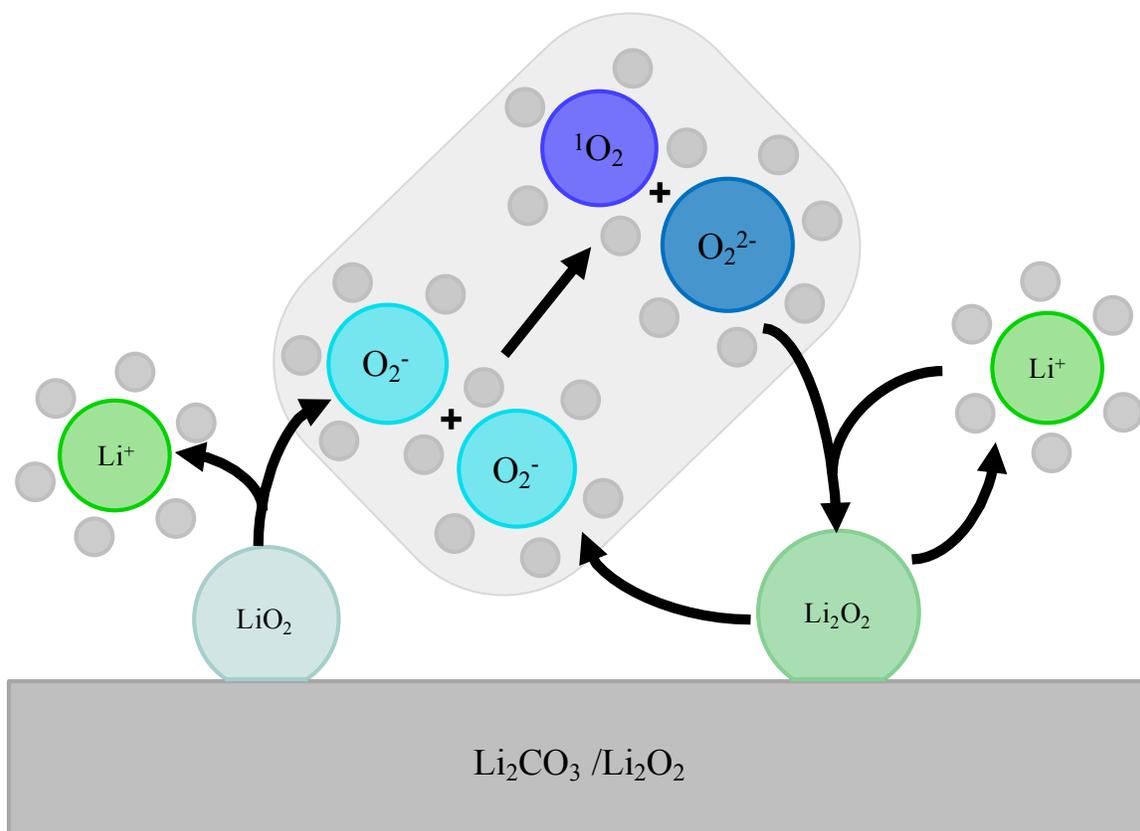


Figure 6.5: Schematic for the process of singlet oxygen production in Li-air and Li-ion batteries. Only the free, versus cation associated, superoxide disproportionation step is shown as that this key reaction in the generation of singlet oxygen. [Reprinted with permission from [177]. Copyright 2020 American Chemical Society]

6.5 Summary

Within this chapter, we have presented a unified picture of the evolution of highly reactive singlet oxygen. Using Marcus theory of kinetics for the disproportionation reactions, we provided insight into how there is a kinetic driving force for the production of the singlet excited state from the reaction of free superoxide. Therefore the interplay of free superoxide interplay with cation associated superoxide controls the proportion of singlet released. Through estimations of the solvation properties of electrolytes as a function of donor and acceptor number based on first-principles calculations, we illuminated the effect of solvent choice on singlet evolution. From this analysis, we were able to rationalize the experimental observation of Li-air, Na-air, and residual Li_2CO_3 in Ni-rich NMC. We proposed a potential-dependent reconstruction of the surface to set the onset potential for the dissolution of superoxide from the surface to the electrolyte and therefore the onset of singlet production. This surface reconstruction is supported by previous first-principles calculations as well as new calculations for the decomposition of Li_2CO_3 presented here.

Chapter 7

Conclusion

The main contributions of this thesis were the development and implementation of a set of tools to increase the efficiency, speed, and accuracy of predictions of battery cathodes as well as to improve upon the understanding of uncertainty in these predictions.

The development of prediction confidence in this thesis enhances the reliability and interpretability of predictions from density functional theory. This was seen in its extension to the search for transition metal ordering within the difficult ternary NMC phase space. This prediction confidence aided in understanding the difference between the predicted stable compositions and the experimentally seen phases and added believably to the predictions through the ability to assign a numerical measure of confidence. This method of prediction confidence has already been adopted to many other problems in the energy storage and conversion further demonstrating its usefulness.

The propagation of uncertainty of the energies fed into a Debye-Grüneisen Model treatment of lattice vibrations allowed for a thorough assessment error. We were able to understand that for many properties, the choice of functional had fortunately little effect on the resulting predictions. We were able to see that properties that relied on higher derivatives of the energy have much larger sensitivity to the input error. This understanding of uncertainty proved of use in assessing the quality of the machine learning potential trained by demonstrating that a negligible level of model error was introduced when extracting thermodynamic properties.

The contributions in employing machine learning represent an advancement in the capabilities of first principles battery modeling by enabling fast optimization that was previously limited by computational cost. This was demonstrated through the prediction of open-circuit voltage, thermodynamic properties, and recreation of structural changes during cycling. The use case in predicting the degree of lattice contraction as a function of various NMC compositions further demonstrates the utility of this machine learning potential.

Next-generation machine learning techniques were discussed and the covariant

convolutional neural network architecture was highlighted.

We also leveraged analytical models of electron transfer kinetics to study disproportionation reactions in batteries. We provide the key insight that free superoxide rather than cation associated superoxide is responsible for the singlet production. The effect of electrolyte properties was explored and a mechanism of potential dependence surface reconstruction was proposed that can explain the onset potential for all experimental observations of singlet oxygen in battery settings.

Chapter 8

Outlook

8.1 Transition Metal Cathodes

The field of transition metal cathodes is increasingly moving towards higher Ni-content. Driven by its increased initial capacity and the concerns around Co, much work is being done to understand and prevent the capacity fading of these materials. The most promising extension of this work from NMC is the inclusion of Al. In layered cathodes, the addition of Al promising as it will increase the intercalation potential, and increases the stability of the structure, suppressing the structural phase transitions seen in LiNiO_2 and high Ni NMC [197]. The addition of Al directly to LiNiO_2 , however, comes at the cost of decreased reversible capacity. The framework presented here, however, presents a road map for how NCA or even NMCA could be investigated. Even more boldly, the presence of Co at all has been recently questioned [198], yet the materials space of Ni-Mn-Al has yet to be fully explored. For even more capacity, the used of lithium-manganese rich materials show promise. These materials have chemical compositions that are a mixture of Li_2MnO_3 and the layered LiMO_2 . Like the Ni-rich layered cathodes, these materials show rapid capacity fade. And as they are Mn-rich, they additionally show an irreversible transformation to the spinel phase. A full computational investigation of these materials is also yet to be performed, presenting another great candidate material space for the methods within this thesis.

8.2 Behler-Parrinello Neural Networks

The work here using the Behler-Parrinello scheme for atom centered symmetry functions present only the beginning of its application to battery research. The full capabilities of the current model have yet to be fully exploited. The ability to accurately reproduce the results of the Debye model predictions from DFT, suggests that a thorough high throughput investigation including the vibrational contribution to

the Gibbs free energy could be performed with a reasonable computational cost. The vibrational properties of even the disordered phases could be studied through the use of special quasirandom structures [199].

A series of data deficiencies were highlighted and the model presented in this thesis could be improved even more with the inclusion of single atom displacements from equilibrium, structure with monoclinic distortions, and even amorphous or structures from high-temperature ab-initio Monte Carlo. Even further improvements would come from the inclusion of cathode surface data. Most of the interesting electrochemistry happens at surfaces and interfaces between the cathode and the electrolyte. An accurate treatment of the cathode a various state of charge and it's explicit interaction with the electrolyte could provide much-needed insight for electrolyte design.

And finally, while we compare much of our results to previous cluster expansion results, we believe there is very much still a place for cluster expansion within a robust computational workflow for battery research. The design loop for the training and convergence cluster expansion is well understood. The addition of the machine learning potential within this workflow, however, presents a supplementation rather than replacement of cluster expansion. The simplicity of the linear fit of cluster expansion means that simulations on the scale of thousands of atoms over hundreds of thousands of Monte Carlo steps can be performed. The course grained results could be passed to the machine learning potential to gain further understanding of thermodynamic properties, more accurate energy, lattice constants, and surface effects. The final predictions could either be validated with DFT and/or through experimentation.

8.3 Beyond Fingerprint Based Learning

Despite the success of the machine learning potential used within this work, there is much promise in the possibility of a more systematic training and hyperparameter selection, as well as a more sophisticated fingerprinting of the material. The covariant convolutional neural network architecture described in Chapter 5 can now predict the energy and force of crystalline materials with orders of magnitude fewer trained coefficients and with competitive accuracy. The extension of this architecture to the direct prediction of vector and tensor quantities is merely a matter of time. It is clear that machine learning architectures that are not directly fingerprint-based represent a likely future in machine learning in materials science and are the center of many studies.

The question of data generation and availability, however, remains an area for vast improvement. Many other fields of machine learning operate with millions of data points, allowing for easy bench-marking of the accuracy and transferable of the machine learning model. In the field of materials science, the limitation in data availability and relatively high cost of high fidelity data generation requires machine learning models to be benchmarked to one type of performance metric, while used for

another possibly different tasks. This is seen clearly in Chapter 4 of this thesis. The data generated for the machine learning potential consisted of random compositions and ordering of Li over various NMC compositions. The accuracy of the potential was benchmarked to how well the potential could recreate the absolute energy of these materials. The main task this potential was used for, however, was the prediction of the open-circuit voltage. That is, the potential needed to accurately predict the energy differences between Li-vacancy ordering to find the most stable ordering at each composition, and then the energy difference between different lithium compositions. While this task is related to the one on which it was benchmarked, success in energy prediction does not guarantee success in energy differences. Two suggestions can be made to address this. The first is to include the prediction of the task at hand (open-circuit voltage and energy differences) into the cost function of the machine learning training. The second would be to include data from the performance of the task at hand. While the data included in Chapter 4 was very similar to the procedure to generate an open-circuit voltage, this was not the case for the prediction of phonons. In both suggestions, sample procedures for the predictions of these derivative properties are needed using calculations at the full fidelity of the underlying data. As this is expensive, some thought needs to be put into how to specifically decide when and how many of these sample calculations to include, and perhaps, how it can be done within a closed loop.

Bibliography

- [1] George E. Blomgren. “The Development and Future of Lithium Ion Batteries”. In: *J. Electrochem. Soc.* 164.1 (2017), A5019–A5025. DOI: 10.1149/2.0251701jes. URL: <http://jes.ecsdl.org/content/164/1/A5019.abstract>.
- [2] Elsa A. Olivetti et al. “Lithium-Ion Battery Supply Chain Considerations: Analysis of Potential Bottlenecks in Critical Metals”. In: *Joule* 1.2 (2018/05/03 2017), pp. 229–243. DOI: 10.1016/j.joule.2017.08.019. URL: <http://dx.doi.org/10.1016/j.joule.2017.08.019>.
- [3] Hoon-Hee Ryu et al. “Capacity Fading of Ni-Rich Li [Ni_xCo_yMn_{1-x-y}]O₂ (0.6 ≤ *x* ≤ 0.95) Cathodes for High-Energy-Density Lithium-Ion Batteries: Bulk or Surface Degradation?” In: *Chemistry of Materials* 30.3 (2018), pp. 1155–1163.
- [4] Johannes Wandt et al. “Singlet oxygen evolution from layered transition metal oxide cathode materials and its implications for lithium-ion batteries”. In: *Mater. Today* 21.8 (2018), pp. 825–833.
- [5] Pierre Hohenberg and Walter Kohn. “Inhomogeneous electron gas”. In: *Phys. Rev.* 136.3B (1964), B864.
- [6] DR Hartree. “The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part I. Theory and Methods”. In: *Math. Proc. Camb. Philos. Soc.* Vol. 24. 1. Cambridge University Press. 1928, pp. 89–110.
- [7] Walter Kohn and Lu Jeu Sham. “Self-consistent equations including exchange and correlation effects”. In: *Physical review* 140.4A (1965), A1133.
- [8] J C Stoddart and P Hanks. “The exchange and correlation energy of a spin-polarised electron gas”. In: *J. Phys. C: Solid State Phys.* 10.16 (1977), p. 3167. URL: <http://stacks.iop.org/0022-3719/10/i=16/a=025>.
- [9] John P Perdew and Alex Zunger. “Self-interaction correction to density-functional approximations for many-electron systems”. In: *Physical Review B* 23.10 (1981), p. 5048.
- [10] David M Ceperley and Berni J Alder. “Ground state of the electron gas by a stochastic method”. In: *Physical Review Letters* 45.7 (1980), p. 566.

- [11] Seymour H Vosko, Leslie Wilk, and Marwan Nusair. “Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis”. In: *Canadian Journal of physics* 58.8 (1980), pp. 1200–1211.
- [12] John P Perdew and Yue Wang. “Accurate and simple analytic representation of the electron-gas correlation energy”. In: *Physical review B* 45.23 (1992), p. 13244.
- [13] Walter Kohn. “Nobel Lecture: Electronic structure of matter—wave functions and density functionals”. In: *Reviews of Modern Physics* 71.5 (1999), p. 1253.
- [14] John P Perdew, Kieron Burke, and Matthias Ernzerhof. “Generalized gradient approximation made simple”. In: *Physical review letters* 77.18 (1996), p. 3865.
- [15] Bjørk Hammer, Lars Bruno Hansen, and Jens Kehlet Nørskov. “Improved adsorption energetics within density-functional theory using revised Perdew-Burke-Ernzerhof functionals”. In: *Physical review B* 59.11 (1999), p. 7413.
- [16] Yue Wang and John P Perdew. “Correlation hole of the spin-polarized electron gas, with exact small-wave-vector and high-density scaling”. In: *Physical Review B* 44.24 (1991), p. 13298.
- [17] Max Dion et al. “Van der Waals density functional for general geometries”. In: *Physical Rev. Lett.* 92.24 (2004), p. 246401.
- [18] Kyuho Lee et al. “Higher-accuracy van der Waals density functional”. In: *Physical Review B* 82.8 (2010), p. 081101.
- [19] Kristian Berland and Per Hyldgaard. “Exchange functional that tests the robustness of the plasmon description of the van der Waals density functional”. In: *Physical Review B* 89.3 (2014), p. 035412.
- [20] Jiri Klimes, David R Bowler, and Angelos Michaelides. “Chemical accuracy for the van der Waals density functional”. In: *Journal of Physics: Condensed Matter* 22.2 (2009), p. 022201.
- [21] Muratahan Aykol, Soo Kim, and C Wolverton. “Van der Waals interactions in layered lithium cobalt oxides”. In: *The Journal of Physical Chemistry C* 119.33 (2015), pp. 19053–19058.
- [22] Anubhav Jain, Yongwoo Shin, and Kristin A Persson. “Computational predictions of energy materials using density functional theory”. In: *Nature Reviews Materials* 1.1 (2016), pp. 1–13.
- [23] Jiawei Zhang et al. “Designing high-performance layered thermoelectric materials through orbital engineering”. In: *Nature Communications* 7 (2016), p. 10892.
- [24] Jeff Greeley et al. “Computational high-throughput screening of electrocatalytic materials for hydrogen evolution”. In: *Nature materials* 5.11 (2006), pp. 909–913.

- [25] Kesong Yang et al. “A search model for topological insulators with high-throughput robustness descriptors”. In: *Nature materials* 11.7 (2012), pp. 614–619.
- [26] Nedko Drebov et al. “Ab initio screening methodology applied to the search for new permanent magnetic materials”. In: *New Journal of Physics* 15.12 (2013), p. 125023.
- [27] Ivano E Castelli et al. “Computational screening of perovskite metal oxides for optimal solar light capture”. In: *Energy & Environmental Science* 5.2 (2012), pp. 5814–5819.
- [28] Yunxing Zuo et al. “Performance and cost assessment of machine learning interatomic potentials”. In: *The Journal of Physical Chemistry A* 124.4 (2020), pp. 731–745.
- [29] Nongnuch Artrith, Alexander Urban, and Gerbrand Ceder. “Efficient and accurate machine-learning interpolation of atomic energies in compositions with many species”. In: *Phys. Rev. B* 96.1 (2017), p. 014112.
- [30] Sönke Lorenz, Axel Groß, and Matthias Scheffler. “Representing high-dimensional potential-energy surfaces for reactions at surfaces by neural networks”. In: *Chemical Physics Letters* 395.4-6 (2004), pp. 210–215.
- [31] Jörg Behler and Michele Parrinello. “Generalized neural-network representation of high-dimensional potential-energy surfaces”. In: *Phys. Rev. Lett.* 98.14 (2007), p. 146401.
- [32] Jörg Behler. “Atom-centered symmetry functions for constructing high-dimensional neural network potentials”. In: *J. Chem. Phys.* 134.7 (2011), p. 074106.
- [33] G Kresse, J Furthmüller, and J Hafner. “Ab initio force constant approach to phonon dispersion relations of diamond and graphite”. In: *EPL (Europhysics Letters)* 32.9 (1995), p. 729.
- [34] Stefano Baroni et al. “Phonons and related crystal properties from density-functional perturbation theory”. In: *Rev. Mod. Phys.* 73 (2 July 2001), pp. 515–562. DOI: 10.1103/RevModPhys.73.515. URL: <https://link.aps.org/doi/10.1103/RevModPhys.73.515>.
- [35] V. L. Moruzzi, J. F. Janak, and K. Schwarz. “Calculated Thermal Properties of Metals”. In: *Phys. Rev. B* 37 (2 Jan. 1988), pp. 790–799. DOI: 10.1103/PhysRevB.37.790. URL: <https://link.aps.org/doi/10.1103/PhysRevB.37.790>.
- [36] Xiao-Gang Lu, Malin Selleby, and Bo Sundman. “Calculations of thermo-physical properties of cubic carbides and nitrides using the Debye–Grüneisen model”. In: *Acta Mater.* 55.4 (2007), pp. 1215–1226. ISSN: 1359-6454. DOI: <https://doi.org/10.1016/j.actamat.2006.05.054>. URL: <http://www.sciencedirect.com/science/article/pii/S1359645406007105>.

- [37] Hu Cui-E et al. “Pressure and Temperature Induced Phase Transition of ZnS from First-Principles Calculations”. In: *Chin. Phys. Lett.* 25.2 (Feb. 2008), pp. 675–678. DOI: 10.1088/0256-307x/25/2/087. URL: <https://doi.org/10.1088/0256-307x/25/2/087>.
- [38] M.A. Blanco et al. “Thermodynamical properties of solids from microscopic theory: applications to MgF_2 and Al_2O_3 ”. In: *J. Mol. Struct. THEOCHEM* 368 (1996). Proceedings of the Second Electronic Computational Chemistry Conference, pp. 245–255. ISSN: 0166-1280. DOI: [https://doi.org/10.1016/S0166-1280\(96\)90571-0](https://doi.org/10.1016/S0166-1280(96)90571-0). URL: <http://www.sciencedirect.com/science/article/pii/S0166128096905710>.
- [39] C. Kittel. *Introduction to Solid State Physics*. John Wiley & Sons, 1996, pp. 117–140.
- [40] O. L. Anderson. “Determination and Some Uses of Isotropic Elastic Constants of Polycrystalline Aggregates Using Single-Crystal Data”. In: *Physical Acoustics Vol. III, Part B: Lattice Dynamics*. Ed. by W. P. Mason. New York: Academic Press, 1965. Chap. 2, pp. 43–97.
- [41] L. D. Landau and E. M. Lifshitz. *A Course on Theoretical Physics Vol. 7, Theory of Elasticity*. London: Pergamon Press, 1970.
- [42] THK Barron. “LXXXII. On the thermal expansion of solids at low temperatures”. In: *The London, Edinburgh, and Dublin Philosophical Magazine and Journal of Science* 46.378 (1955), pp. 720–734.
- [43] Kurt Lejaeghere et al. “Error estimates for solid-state density-functional theory predictions: an overview by means of the ground-state elemental crystals”. In: *Critical Reviews in Solid State and Materials Sciences* 39.1 (2014), pp. 1–24.
- [44] Narbe Mardirossian and Martin Head-Gordon. “Thirty years of density functional theory in computational chemistry: an overview and extensive assessment of 200 density functionals”. In: *Molecular Physics* 115.19 (2017), pp. 2315–2372.
- [45] Pascal Pernot and Andreas Savin. “Probabilistic performance estimators for computational chemistry methods: The empirical cumulative distribution function of absolute errors”. In: *J. Chem. Phys.* 148.24 (2018), p. 241707. DOI: 10.1063/1.5016248. URL: <https://doi.org/10.1063/1.5016248>.
- [46] Sam De Waele et al. “Error estimates for density-functional theory predictions of surface energy and work function”. In: *Phys. Rev. B* 94 (23 Dec. 2016), p. 235418. DOI: 10.1103/PhysRevB.94.235418. URL: <https://link.aps.org/doi/10.1103/PhysRevB.94.235418>.

- [47] Kurt Lejaeghere et al. “Is the Error on First-Principles Volume Predictions Absolute or Relative?” In: *Comput. Mater. Sci.* 117 (2016), pp. 390–396. ISSN: 0927-0256. DOI: <https://doi.org/10.1016/j.commatsci.2016.01.039>. URL: <http://www.sciencedirect.com/science/article/pii/S092702561630012X>.
- [48] Søren L Frederiksen et al. “Bayesian ensemble approach to error estimation of interatomic potentials”. In: *Physical review letters* 93.16 (2004), p. 165501.
- [49] Jens Jørgen Mortensen et al. “Bayesian error estimation in density-functional theory”. In: *Phys. Rev. Lett.* 95.21 (2005), p. 216401.
- [50] V Petzold, T Bligaard, and KW Jacobsen. “Construction of new electronic density functionals with error estimation through fitting”. In: *Topics in Catalysis* 55.5-6 (2012), pp. 402–417.
- [51] Jess Wellendorff et al. “Density functionals for surface science: Exchange-correlation model development with Bayesian error estimation”. In: *Phys. Rev. B* 85 (23 June 2012), p. 235149.
- [52] Jess Wellendorff et al. “mBEEF: An accurate semi-local Bayesian error estimation density functional”. In: *J. of Chem. Phys.* 140.14 (2014), p. 144107.
- [53] Pascal Pernot. “The parameter uncertainty inflation fallacy”. In: *The Journal of Chemical Physics* 147.10 (2017), p. 104102.
- [54] Jonny Proppe et al. “Uncertainty quantification for quantum chemical models of complex reaction networks”. In: *Faraday Discussions* 195 (2017), pp. 497–520.
- [55] Manuel Aldegunde, James R Kermode, and Nicholas Zabaras. “Development of an exchange–correlation functional with uncertainty quantification capabilities for density functional theory”. In: *Journal of Computational Physics* 311 (2016), pp. 173–195.
- [56] Gregor N Simm and Markus Reiher. “Systematic error estimation for chemical reaction energies”. In: *Journal of chemical theory and computation* 12.6 (2016), pp. 2762–2773.
- [57] Zeeshan Ahmad and Venkatasubramanian Viswanathan. “Quantification of uncertainty in first-principles predicted mechanical properties of solids: Application to solid ion conductors”. In: *Phys. Rev. B* 94 (6 Aug. 2016), p. 064105. DOI: 10.1103/PhysRevB.94.064105. URL: <https://link.aps.org/doi/10.1103/PhysRevB.94.064105>.
- [58] Gregory Houchins and Venkatasubramanian Viswanathan. “Quantifying confidence in density functional theory predictions of magnetic ground states”. In: *Phys. Rev. B* 96 (13 Oct. 2017), p. 134426. DOI: 10.1103/PhysRevB.96.134426. URL: <https://link.aps.org/doi/10.1103/PhysRevB.96.134426>.

- [59] Pin-Wen Guan, Gregory Houchins, and Venkatasubramanian Viswanathan. “Uncertainty quantification of DFT-predicted finite temperature thermodynamic properties within the Debye model”. In: *J. Chem. Phys.* 151.24 (2019), p. 244702.
- [60] Holden L Parks, Alan JH McGaughey, and Venkatasubramanian Viswanathan. “Uncertainty quantification in first-principles predictions of harmonic vibrational frequencies of molecules and molecular complexes”. In: *The Journal of Physical Chemistry C* 123.7 (2019), pp. 4072–4084.
- [61] Holden L Parks et al. “Uncertainty quantification in first-principles predictions of phonon properties and lattice thermal conductivity”. In: *arXiv preprint arXiv:2005.08835* (2020).
- [62] Rune Christensen, Heine A. Hansen, and Tejs Vegge. “Identifying systematic DFT errors in catalytic reactions”. In: *Catal. Sci. Technol.* 5 (11 2015), pp. 4946–4949. DOI: 10.1039/C5CY01332A. URL: <http://dx.doi.org/10.1039/C5CY01332A>.
- [63] Rune Christensen et al. “Functional Independent Scaling Relation for ORR/OER Catalysts”. In: *J. Phys. Chem. C* 120.43 (2016), pp. 24910–24916. DOI: 10.1021/acs.jpcc.6b09141.
- [64] Wenhao Sun et al. “The thermodynamic scale of inorganic crystalline metastability”. In: *Science Advances* 2.11 (2016).
- [65] S. H. Vosko, J. P. Perdew, and A. H. MacDonald. “Ab Initio Calculation of the Spin Susceptibility for the Alkali Metals Using the Density-Functional Formalism”. In: *Phys. Rev. Lett.* 35 (25 Dec. 1975), pp. 1725–1728. DOI: 10.1103/PhysRevLett.35.1725. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.35.1725>.
- [66] U K Poulsen, J Kollar, and O K Andersen. “Magnetic and cohesive properties from canonical bands (for transition metals)”. In: *J. Phys. F: Met. Phys.* 6.9 (1976), p. L241. URL: <http://stacks.iop.org/0305-4608/6/i=9/a=002>.
- [67] O. Gunnarsson. “Band model for magnetism of transition metals in the spin-density-functional formalism”. In: *J. Phys. F: Met. Phys.* 6 (Apr. 1976), pp. 587–606. DOI: 10.1088/0305-4608/6/4/018.
- [68] O. Gunnarsson. “Band magnetism in the spin-density-functional formalism”. In: *J. Appl. Phys.* 49 (Mar. 1978), pp. 1399–1404. DOI: 10.1063/1.325005.
- [69] R M Bozorth. “Magnetic Properties of Materials”. In: *American Institute of Physics Handbook*. Ed. by D E Gray. 1957. Chap. 5.
- [70] Eric Fawcett. “Spin-density-wave antiferromagnetism in chromium”. In: *Rev. Mod. Phys.* 60 (1 Jan. 1988), pp. 209–283.
- [71] P. D. Battle et al. “A novel magnetic phase transition in anhydrous iron (III) phosphate, FePO₄”. In: *J. Phys. C: Solid State Phys.* 15.26 (1982), p. L919.

- [72] G. Rousse et al. “Magnetic Structures of the Triphylite LiFePO_4 and of Its Delithiated Form FePO_4 ”. In: *Chem. Mater.* 15.21 (2003), pp. 4082–4090.
- [73] J. Rodriguez-Carvajal et al. “Neutron-diffraction study of the Jahn-Teller transition in stoichiometric LaMnO_3 ”. In: *Phys. Rev. B* 57 (6 Feb. 1998), R3189–R3192.
- [74] R. Hafner et al. “Magnetic ground state of Cr in density-functional theory”. In: *Phys. Rev. B* 65 (18 May 2002), p. 184432. DOI: 10.1103/PhysRevB.65.184432. URL: <https://link.aps.org/doi/10.1103/PhysRevB.65.184432>.
- [75] H Sawada et al. “Jahn-Teller distortion and magnetic structures in LaMnO_3 ”. In: *Phys. Rev. B* 56.19 (1997), p. 12154.
- [76] Matthew R. Suchomel et al. “Spin-induced symmetry breaking in orbitally ordered NiCr_2O_4 and CuCr_2O_4 ”. In: *Phys. Rev. B* 86 (5 Aug. 2012), p. 054406. DOI: 10.1103/PhysRevB.86.054406. URL: <https://link.aps.org/doi/10.1103/PhysRevB.86.054406>.
- [77] John H Van Vleck. “A Survey of the Theory of Ferromagnetism”. In: *Rev. Mod. Phys.* 17.1 (1945), p. 27.
- [78] K. Capelle and L. N. Oliveira. “Density-functional theory for spin-density waves and antiferromagnetic systems”. In: *Phys. Rev. B* 61 (22 June 2000), pp. 15228–15240. DOI: 10.1103/PhysRevB.61.15228. URL: <https://link.aps.org/doi/10.1103/PhysRevB.61.15228>.
- [79] E A Owen and G I Williams. “A low-temperature X-ray camera”. In: *J. Sci. Instrum.* 31.2 (1954), p. 49.
- [80] Ralph Wyckoff and Walter Graystone. *Crystal structures*. Malabar, FL: Krieger, 1964.
- [81] Gary J. Long, Anthony K. Cheetham, and Peter D. Battle. “Study of the iron-phosphorus-oxygen system by Moessbauer effect, neutron diffraction, magnetic susceptibility, and analytical electron microscopy: some pitfalls and solutions in the analysis of a complex mixture”. In: *Inorg. Chem.* 22.21 (1983), pp. 3012–3016.
- [82] E. Prince. “Crystal and Magnetic Structure of Copper Chromite”. In: *Acta Crystallographica* 10.9 (Sept. 1957), pp. 554–556.
- [83] S. Lebègue M. Ramzan and R. Ahuja. “Ab initio study of lithium and sodium iron fluorophosphate cathodes for rechargeable batteries”. In: *Appl. Phys. Lett.* 94.15 (2009), p. 151904. DOI: 10.1063/1.3119704. URL: <http://dx.doi.org/10.1063/1.3119704>.

- [84] Fei Zhou, Thomas Maxisch, and Gerbrand Ceder. “Configurational Electronic Entropy and the Phase Diagram of Mixed-Valence Oxides: The Case of Li_xFePO_4 ”. In: *Phys. Rev. Lett.* 97 (15 Oct. 2006), p. 155704. DOI: 10.1103/PhysRevLett.97.155704. URL: <https://link.aps.org/doi/10.1103/PhysRevLett.97.155704>.
- [85] Dilip Krishnamurthy, Vaidish Sumaria, and Venkatasubramanian Viswanathan. “Quantifying robustness of DFT predicted pathways and activity determining elementary steps for electrochemical reactions”. In: *The Journal of Chemical Physics* 150.4 (2019), p. 041717.
- [86] Olga Vinogradova et al. “Quantifying confidence in DFT-predicted surface pourbaix diagrams of transition-metal electrode–electrolyte interfaces”. In: *Langmuir* 34.41 (2018), pp. 12259–12269.
- [87] Victor Venturi et al. “Machine Learning Enabled Discovery of Application Dependent Design Principles for Two-dimensional Materials”. In: *arXiv preprint arXiv:2003.13418* (2020).
- [88] Vikram Pande and Venkatasubramanian Viswanathan. “Robust high-fidelity DFT study of the lithium-graphite phase diagram”. In: *Physical Review Materials* 2.12 (2018), p. 125401.
- [89] Ying Yuan et al. “Uncertainty Quantification of First Principles Computational Phase Diagram Predictions of Li-Si System Via Bayesian Sampling”. In: *arXiv preprint arXiv:2003.13393* (2020).
- [90] Peter J Feibelman et al. “The CO/Pt (111) Puzzle”. In: *The Journal of Physical Chemistry B* 105.18 (2001), pp. 4018–4025.
- [91] Patanachai Janthon et al. “Adding pieces to the CO/Pt (111) puzzle: the role of dispersion”. In: *The Journal of Physical Chemistry C* 121.7 (2017), pp. 3970–3977.
- [92] Abhirup Patra et al. “Rethinking CO adsorption on transition-metal surfaces: Effect of density-driven self-interaction errors”. In: *Physical Review B* 100.3 (2019), p. 035442.
- [93] Noah H. Paulson, Elise Jennings, and Marius Stan. “Bayesian strategies for uncertainty quantification of the thermodynamic properties of materials”. In: *Int. J. Eng. Sci.* 142 (2019), pp. 74–93. ISSN: 0020-7225. DOI: <https://doi.org/10.1016/j.ijengsci.2019.05.011>. URL: <http://www.sciencedirect.com/science/article/pii/S0020722518314721>.
- [94] Pascal Vinet et al. “Temperature effects on the universal equation of state of solids”. In: *Phys. Rev. B* 35 (4 Feb. 1987), pp. 1945–1953. DOI: 10.1103/PhysRevB.35.1945. URL: <https://link.aps.org/doi/10.1103/PhysRevB.35.1945>.

- [95] William H Evans et al. “Thermodynamic properties of the alkali metals”. In: *J. Res. Natl. Bur. Stand. (U. S.)* 55.2 (1955), pp. 83–96.
- [96] DW Jeppson et al. *Lithium literature review: lithium’s properties and interactions*. Hanford Engineering Development Lab., 1978. DOI: 10.2172/6885395.
- [97] Robert Weast. *CRC handbook of chemistry and physics*. Cleveland, Ohio: CRC Press”, 1977.
- [98] G.V. Samsonov. *Handbook of the Physicochemical Properties of the Elements*. Springer US, 2012. ISBN: 9781468460667. URL: <https://books.google.com/books?id=ZzzjBwAAQBAJ>.
- [99] D.E. Gray, ed. *American Institute of Physics Handbook*. McGraw-Hill handbooks. McGraw-Hill, 1963. URL: <https://books.google.com/books?id=IxxIAQAATAAJ>.
- [100] H Kawaji et al. “Low temperature heat capacity and thermodynamic functions of LiCoO₂”. In: *Journal of thermal analysis and calorimetry* 68.3 (2002), pp. 833–839.
- [101] Petronela Gotcu-Freis et al. “New experimental heat capacity and enthalpy of formation of lithium cobalt oxide”. In: *The Journal of Chemical Thermodynamics* 84 (2015), pp. 118–127.
- [102] Eric Jianfeng Cheng et al. “Elastic properties of lithium cobalt oxide (LiCoO₂)”. In: *Journal of Asian Ceramic Societies* 5.2 (2017), pp. 113–117.
- [103] M Stanley Whittingham. “Electrical energy storage and intercalation chemistry”. In: *Science* 192.4244 (1976), pp. 1126–1127.
- [104] K Mizushima et al. “Li_xCoO₂ (0 ≤ x ≤ 1): A new cathode material for batteries of high energy density”. In: *Materials Research Bulletin* 15.6 (1980), pp. 783–789.
- [105] Jaephil Cho et al. “Zero-strain intercalation cathode for rechargeable Li-Ion cell”. In: *Angewandte Chemie* 113.18 (2001), pp. 3471–3473.
- [106] Zhaohui Chen and JR Dahn. “Methods to obtain excellent capacity retention in LiCoO₂ cycled to 4.5 V”. In: *Electrochimica Acta* 49.7 (2004), pp. 1079–1090.
- [107] Jaephil Cho, Yong Jeong Kim, and Byungwoo Park. “Novel LiCoO₂ cathode material with Al₂O₃ coating for a Li ion cell”. In: *Chemistry of Materials* 12.12 (2000), pp. 3788–3791.
- [108] JR Dahn et al. “Rechargeable LiNiO₂/carbon cells”. In: *Journal of the Electrochemical Society* 138.8 (1991), p. 2207.
- [109] Wu Li, JN Reimers, and JR Dahn. “In situ x-ray diffraction and electrochemical studies of Li_{1-x}NiO₂”. In: *Solid State Ionics* 67.1-2 (1993), pp. 123–130.

- [110] Hongyang Li et al. “Updating the Structure and Electrochemistry of Li_xNiO_2 for $0 \leq x \leq 1$ ”. In: *Journal of The Electrochemical Society* 165.13 (2018), A2985.
- [111] G. Vitins and K. West. “Lithium Intercalation into Layered LiMnO_2 ”. In: *J. Electrochem. Soc.* 144.8 (1997), pp. 2587–2592. DOI: 10.1149/1.1837869. URL: <http://jes.ecsdl.org/content/144/8/2587.abstract>.
- [112] AE Abdel-Ghany et al. “Structural properties and electrochemistry of α - LiFeO_2 ”. In: *Journal of Power Sources* 197 (2012), pp. 285–291.
- [113] John Reed and Gerbrand Ceder. “Role of electronic structure in the susceptibility of metastable transition-metal oxide structures to transformation”. In: *Chemical reviews* 104.10 (2004), pp. 4513–4534.
- [114] Zhaolin Liu, Aishui Yu, and Jim Y Lee. “Synthesis and characterization of $\text{LiNi}_{1-x}\text{Co}_x\text{Mn}_y\text{O}_2$ as the cathode materials of secondary lithium batteries”. In: *Journal of Power Sources* 81 (1999), pp. 416–419.
- [115] Tsutomu Ohzuku and Yoshinari Makimura. “Layered lithium insertion material of $\text{LiCo}_{1/3}\text{Ni}_{1/3}\text{Mn}_{1/3}\text{O}_2$ for lithium-ion batteries”. In: *Chemistry Letters* 30.7 (2001), pp. 642–643.
- [116] Naoaki Yabuuchi and Tsutomu Ohzuku. “Novel lithium insertion material of $\text{LiCo}_{1/3}\text{Ni}_{1/3}\text{Mn}_{1/3}\text{O}_2$ for advanced lithium-ion batteries”. In: *Journal of Power Sources* 119 (2003), pp. 171–174.
- [117] Seong-Min Bak et al. “Structural changes and thermal stability of charged $\text{LiNi}_x\text{Mn}_y\text{Co}_z\text{O}_2$ cathode materials studied by combined in situ time-resolved XRD and mass spectroscopy”. In: *ACS applied materials & interfaces* 6.24 (2014), pp. 22594–22601.
- [118] Maryam Masoumi et al. “Enthalpies of formation of layered $\text{LiNi}_x\text{Mn}_x\text{Co}_{1-2x}\text{O}_2$ ($0 \leq x \leq 0.5$) compounds as lithium ion battery cathode materials”. In: *International Journal of Materials Research* 108.11 (2017), pp. 869–878.
- [119] Dongli Zeng et al. “Cation Ordering in $\text{Li}[\text{Ni}_x\text{Mn}_x\text{Co}_{(1-2x)}]\text{O}_2$ -Layered Cathode Materials: A Nuclear Magnetic Resonance (NMR), Pair Distribution Function, X-ray Absorption Spectroscopy, and Electrochemical Study”. In: *Chemistry of Materials* 19.25 (2007), pp. 6277–6289.
- [120] Feng Lin et al. “Surface reconstruction and chemical evolution of stoichiometric layered cathode materials for lithium-ion batteries”. In: *Nature communications* 5.1 (2014), pp. 1–9.
- [121] Rahul Malik, Fei Zhou, and Gerbrand Ceder. “Phase diagram and electrochemical properties of mixed olivines from first-principles calculations”. In: *Phys. Rev. B* 79 (21 June 2009), p. 214201. DOI: 10.1103/PhysRevB.79.214201. URL: <https://link.aps.org/doi/10.1103/PhysRevB.79.214201>.

- [122] Eunseok Lee, Hakim Iddir, and Roy Benedek. “Rapidly convergent cluster expansion and application to lithium ion battery materials”. In: *Phys. Rev. B* 95 (8 Feb. 2017), p. 085134. DOI: 10.1103/PhysRevB.95.085134. URL: <https://link.aps.org/doi/10.1103/PhysRevB.95.085134>.
- [123] Anton Van der Ven et al. “First-principles investigation of phase stability in Li_xCoO_2 ”. In: *Phy. Rev. B* 58.6 (1998), p. 2975.
- [124] C Wolverton and Alex Zunger. “First-principles prediction of vacancy order-disorder and intercalation battery voltages in Li_xCoO_2 ”. In: *Phys. Rev. Lett.* 81.3 (1998), p. 606.
- [125] A Van der Ven and G Ceder. “Ordering in $\text{Li}_x(\text{Ni}_{0.5}\text{Mn}_{0.5})\text{O}_2$ and its relation to charge capacity and electrochemical behavior in rechargeable lithium batteries”. In: *Electrochem. Comm.* 6.10 (2004), pp. 1045–1050.
- [126] Gregory Houchins. *Layered Oxide Cluster Expansion*. <https://gitlab.com/ghouchin/layered-oxide-cluster-expansion>. 2020.
- [127] Alim B. Alchagirov et al. “Energy and pressure versus volume: Equations of state motivated by the stabilized jellium model”. In: *Phys. Rev. B* 63 (22 May 2001), p. 224115.
- [128] H. Akaike. “A new look at the statistical model identification”. In: *IEEE Transactions on Automatic Control* 19.6 (Dec. 1974), pp. 716–723. ISSN: 0018-9286. DOI: 10.1109/TAC.1974.1100705.
- [129] Mervyn Stone. “An asymptotic equivalence of choice of model by cross-validation and Akaike’s criterion”. In: *Journal of the Royal Statistical Society: Series B (Methodological)* 39.1 (1977), pp. 44–47.
- [130] Clifford M. Hurvich and Chih-Ling Tsai. “Regression and Time Series Model Selection in Small Samples”. In: *Biometrika* 76.2 (1989), pp. 297–307. ISSN: 00063444. URL: <http://www.jstor.org/stable/2336663>.
- [131] Gideon Schwarz. “Estimating the dimension of a model”. In: *Ann. Stat* 6.2 (1978), pp. 461–464.
- [132] Gregory Houchins and Venkatasubramanian Viswanathan. “Towards Ultra Low Cobalt Cathodes: A High Fidelity Computational Phase Search of Layered Li-Ni-Mn-Co Oxides”. In: *J. Electrochem. Soc.* 167.7 (2019), p. 070506.
- [133] Colby R. Brown et al. “Combinatorial study of the Li-Ni-Mn-Co oxide pseudoquaternary system for use in Li-ion battery materials research”. In: *ACS Comb. Sci.* 17 (2015), pp. 381–391. ISSN: 21568952.
- [134] Natasha A. Chernova et al. “Layered $\text{Li}_x\text{Ni}_y\text{Mn}_y\text{Co}_{1-2y}\text{O}_2$ Cathodes for Lithium Ion Batteries: Understanding Local Structure via Magnetic Properties”. In: *Chem. Mater.* 19.19 (Sept. 2007), pp. 4682–4693.

- [135] Christian Julien et al. “Optimization of Layered Cathode Materials for Lithium-Ion Batteries”. In: *Materials* 9.7 (2016), p. 595. ISSN: 1996-1944. DOI: 10.3390/ma9070595. URL: <http://www.mdpi.com/1996-1944/9/7/595>.
- [136] Gregory Houchins and Venkatasubramanian Viswanathan. “An accurate machine-learning calculator for optimization of Li-ion battery cathodes”. In: *The Journal of Chemical Physics* 153.5 (2020), p. 054124.
- [137] Kirthevasan Kandasamy et al. “Tuning hyperparameters without grad students: Scalable and robust Bayesian optimisation with Dragonfly”. In: *arXiv preprint arXiv:1903.06694* (2019).
- [138] Peter Auer. “Using confidence bounds for exploitation-exploration trade-offs”. In: *J. Mach. Learn. Res.* 3.Nov (2002), pp. 397–422.
- [139] Niranjana Srinivas et al. “Gaussian process optimization in the bandit setting: no regret and experimental design”. In: *Proceedings of the 27th International Conference on International Conference on Machine Learning*. 2010, pp. 1015–1022.
- [140] William R Thompson. “On the likelihood that one unknown probability exceeds another in view of the evidence of two samples”. In: *Biometrika* 25.3/4 (1933), pp. 285–294.
- [141] Donald R Jones, Matthias Schonlau, and William J Welch. “Efficient global optimization of expensive black-box functions”. In: *J. Global Optim.* 13.4 (1998), pp. 455–492.
- [142] Chao Qin, Diego Klabjan, and Daniel Russo. “Improving the expected improvement algorithm”. In: *Advances in Neural Information Processing Systems*. 2017, pp. 5381–5391.
- [143] A Togo and I Tanaka. “First principles phonon calculations in materials science”. In: *Scr. Mater.* 108 (Nov. 2015), pp. 1–5.
- [144] D. Frenkel. “Free-energy computations and first-order phase transitions”. In: *Molecular-dynamics Simulation of Statistical-Mechanical Systems*. Ed. by Giovanni Ciccotti and William G. Hoover. Amsterdam, 1985, p. 151.
- [145] Anne Boutin, Bernard Tavitian, Alain H Fuchs, et al. “Grand canonical Monte Carlo simulations of adsorption of mixtures of xylene molecules in faujasite zeolites”. In: *Faraday Discuss.* 106 (1997), pp. 307–323.
- [146] A van Bondi. “van der Waals volumes and radii”. In: *J. Phys. Chem.* 68.3 (1964), pp. 441–451.
- [147] Malcolm W Chase Jr. “NIST-JANAF thermochemical tables”. In: *J. Phys. Chem. Ref. Data, Monograph* 9 (1998).

- [148] ME Arroyo y de Dompablo, A Van der Ven, and G Ceder. “First-principles calculations of lithium ordering and phase stability on Li_xNiO_2 ”. In: *Physical Review B* 66.6 (2002), p. 064112.
- [149] H Xia et al. “Electrochemical Behavior and Li Diffusion Study of LiCoO_2 Thin Film Electrodes Prepared by PLD”. In: (2007).
- [150] Keke Chang, Bengt Hallstedt, and Denis Music. “Thermodynamic description of the LiNiO_2 – NiO_2 pseudo-binary system and extrapolation to the $\text{Li}(\text{Co}, \text{Ni})\text{O}_2$ – $(\text{Co}, \text{Ni})\text{O}_2$ system”. In: *Calphad* 37 (2012), pp. 100–107.
- [151] ME Arroyo y de Dompablo and G Ceder. “On the Origin of the Monoclinic Distortion in Li_xNiO_2 ”. In: *Chemistry of materials* 15.1 (2003), pp. 63–67.
- [152] Florian Schipper et al. “Recent advances and remaining challenges for lithium ion battery cathodes”. In: *Journal of the Electrochemical Society* 164.1 (2016), A6220.
- [153] Katharina Märker et al. “Evolution of structure and lithium dynamics in $\text{LiNi}_{0.8}\text{Mn}_{0.1}\text{Co}_{0.1}\text{O}_2$ (NMC811) cathodes during electrochemical cycling”. In: *Chemistry of Materials* 31.7 (2019), pp. 2545–2554.
- [154] Justin S Smith, Olexandr Isayev, and Adrian E Roitberg. “ANI-1: an extensible neural network potential with DFT accuracy at force field computational cost”. In: *Chemical science* 8.4 (2017), pp. 3192–3203.
- [155] Felix A Faber et al. “Alchemical and structural distribution based representation for universal quantum machine learning”. In: *The Journal of Chemical Physics* 148.24 (2018), p. 241717.
- [156] Albert P Bartók et al. “Gaussian approximation potentials: The accuracy of quantum mechanics, without the electrons”. In: *Physical review letters* 104.13 (2010), p. 136403.
- [157] Aidan P Thompson et al. “Spectral neighbor analysis method for automated generation of quantum-accurate interatomic potentials”. In: *Journal of Computational Physics* 285 (2015), pp. 316–330.
- [158] Albert P Bartók, Risi Kondor, and Gábor Csányi. “On representing chemical environments”. In: *Physical Review B* 87.18 (2013), p. 184115.
- [159] Matthias Rupp et al. “Fast and accurate modeling of molecular atomization energies with machine learning”. In: *Physical review letters* 108.5 (2012), p. 058301.
- [160] Li Zhu et al. “A fingerprint based metric for measuring similarities of crystalline structures”. In: *The Journal of chemical physics* 144.3 (2016), p. 034203.
- [161] Tian Xie and Jeffrey C Grossman. “Crystal graph convolutional neural networks for an accurate and interpretable prediction of material properties”. In: *Physical review letters* 120.14 (2018), p. 145301.

- [162] Han Wang et al. “DeePMD-kit: A deep learning package for many-body potential energy representation and molecular dynamics”. In: *Computer Physics Communications* 228 (2018), pp. 178–184.
- [163] Kristof T Schütt et al. “Quantum-chemical insights from deep tensor neural networks”. In: *Nature communications* 8.1 (2017), pp. 1–8.
- [164] Stefan Chmiela et al. “Machine learning of accurate energy-conserving molecular force fields”. In: *Science advances* 3.5 (2017), e1603015.
- [165] Kristof Schütt et al. “SchNet: A continuous-filter convolutional neural network for modeling quantum interactions”. In: *Advances in neural information processing systems*. 2017, pp. 991–1001.
- [166] Brandon Anderson, Truong Son Hy, and Risi Kondor. “Cormorant: Covariant molecular neural networks”. In: *Advances in Neural Information Processing Systems*. 2019, pp. 14537–14546.
- [167] Risi Kondor, Zhen Lin, and Shubhendu Trivedi. “Clebsch–gordan nets: a fully fourier space spherical convolutional neural network”. In: *Advances in Neural Information Processing Systems*. 2018, pp. 10117–10126.
- [168] Taco S. Cohen et al. “Spherical CNNs”. In: *CoRR* abs/1801.10130 (2018). arXiv: 1801.10130. URL: <http://arxiv.org/abs/1801.10130>.
- [169] Truong Son Hy et al. “Predicting molecular properties with covariant compositional networks”. In: *The Journal of chemical physics* 148.24 (2018), p. 241745.
- [170] L. D. Landau and E. M. Lifshitz. *A Course on Theoretical Physics Vol. 3, Quantum Mechanics; Non-relativistic Theory*. London: Pergamon Press, 1965.
- [171] Claude Schweitzer and Reinhard Schmidt. “Physical mechanisms of generation and deactivation of singlet oxygen”. In: *Chem. Rev.* 103.5 (2003), pp. 1685–1758.
- [172] Ahsan U Khan. “Singlet molecular oxygen from superoxide anion and sensitized fluorescence of organic molecules”. In: *Science* 168.3930 (1970), pp. 476–477.
- [173] Edward A Mayeda and Allen J Bard. “Production of singlet oxygen in electrogenerated radical ion electron transfer reactions”. In: *J. Am. Chem. Soc.* 95.19 (1973), pp. 6223–6226.
- [174] Ahsan U Khan. “Direct spectral evidence of the generation of singlet molecular oxygen ($^1\Delta_g$) in the reaction of potassium superoxide with water”. In: *J. Am. Chem. Soc.* 103.21 (1981), pp. 6516–6517.
- [175] EJ Corey, Mukund M Mehrotra, and Ahsan U Khan. “Water induced dismutation of superoxide anion generates singlet molecular oxygen”. In: *Biochem. Biophys. Res. Commun.* 145.2 (1987), pp. 842–846.

- [176] Rudolph A Marcus. “Electron transfer reactions in chemistry. Theory and experiment”. In: *Rev. Mod. Phys.* 65.3 (1993), p. 599.
- [177] Gregory Houchins, Vikram Pande, and Venkatasubramanian Viswanathan. “Mechanism for Singlet Oxygen Production in Li-Ion and Metal-Air Batteries”. In: *ACS Energy Letters* (2020).
- [178] Vyacheslav S Bryantsev, Mario Blanco, and Francesco Faglioni. “Stability of lithium superoxide LiO_2 in the gas phase: computational study of dimerization and disproportionation reactions”. In: *J. Phys. Chem. A* 114.31 (2010), pp. 8165–8169.
- [179] Kent M Ervin et al. “The Only Stable State of O_2^- is the $X^2\Pi_g$ Ground State and It (Still!) Has an Adiabatic Electron Detachment Energy of 0.45 eV”. In: *J. Phys. Chem. A* 107.41 (2003), pp. 8521–8529.
- [180] Ahsan U Khan. “Activated oxygen: singlet molecular oxygen and superoxide anion”. In: *Photochem. Photobiol.* 28.4-5 (1978), pp. 615–626.
- [181] Lei Cheng et al. “Computational studies of solubilities of LiO_2 and Li_2O_2 in aprotic solvents”. In: *J. Electrochem. Soc.* 164.11 (2017), E3696–E3701.
- [182] David G Kwabi et al. “Experimental and Computational Analysis of the Solvent-Dependent $\text{O}_2/\text{Li}^+-\text{O}_2^-$ Redox Couple: Standard Potentials, Coupling Strength, and Implications for Lithium–Oxygen Batteries”. In: *Angew. Chem. Int. Ed.* 55.9 (2016), pp. 3129–3134.
- [183] Ahsan U Khan. “Theory of electron transfer generation and quenching of singlet oxygen [$^1\Sigma_g^+$ and $^1\Delta_g$] by superoxide anion. The role of water in the dismutation of superoxide anion”. In: *J. Am. Chem. Soc.* 99.2 (1977), pp. 370–371.
- [184] Eléonore Mourad et al. “Singlet oxygen from cation driven superoxide disproportionation and consequences for aprotic metal– O_2 batteries”. In: *Energy Environ. Sci.* 12.8 (2019), pp. 2559–2568.
- [185] Chuanchao Sheng et al. “Disproportionation of sodium superoxide in metal–air batteries”. In: *Angewandte Chemie* 130.31 (2018), pp. 10054–10058.
- [186] Johannes Wandt et al. “Singlet oxygen formation during the charging process of an aprotic lithium–oxygen battery”. In: *Angew. Chem. Int. Ed.* 55.24 (2016), pp. 6892–6895.
- [187] Nika Mahne et al. “Singlet oxygen generation as a major cause for parasitic reactions during cycling of aprotic lithium–oxygen batteries”. In: *Nat. Energy* 2.5 (2017), p. 17036.
- [188] Alan C Luntz and Bryan D McCloskey. “Li–air batteries: Importance of singlet oxygen”. In: *Nat. Energy* 2.5 (2017), p. 17056.

- [189] JS Hummelshøj, AC Luntz, and JK Nørskov. “Theoretical evidence for low kinetic overpotentials in Li–O₂ electrochemistry”. In: *J. Chem. Phys.* 138.3 (2013), p. 034703.
- [190] Lukas Schafzahl et al. “Singlet oxygen during cycling of the aprotic sodium–O₂ battery”. In: *Angew. Chem. Int. Ed.* 56.49 (2017), pp. 15728–15732.
- [191] Byungju Lee et al. “First-principles study of the reaction mechanism in sodium–oxygen batteries”. In: *Chem. Mater.* 26.2 (2014), pp. 1048–1055.
- [192] Conrad L Bender et al. “One-or Two-Electron Transfer? The Ambiguous Nature of the Discharge Products in Sodium–Oxygen Batteries”. In: *Angew. Chem. Int. Ed.* 55.15 (2016), pp. 4640–4649.
- [193] Roland Jung et al. “Oxygen release and its effect on the cycling stability of LiNi_xMn_yCo_zO₂ (NMC) cathode materials for Li-ion batteries”. In: *J. Electrochem. Soc.* 164.7 (2017), A1361–A1377.
- [194] Sara E Renfrew and Bryan D McCloskey. “Residual lithium carbonate predominantly accounts for first cycle CO₂ and CO outgassing of Li-stoichiometric and Li-rich layered transition-metal oxides”. In: *J. Am. Chem. Soc.* 139.49 (2017), pp. 17853–17860.
- [195] H Arai et al. “Characterization and cathode performance of Li_{1-x}Ni_{1+x}O₂ prepared with the excess lithium method”. In: *Solid State Ionics* 80.3-4 (1995), pp. 261–269.
- [196] Min Yu and Dallas R Trinkle. “Accurate and efficient algorithm for Bader charge integration”. In: *J. Chem. Phys.* 134.6 (2011), p. 064111.
- [197] Jing Xu et al. “A review of Ni-based layered oxides for rechargeable Li-ion batteries”. In: *Journal of Materials Chemistry A* 5.3 (2017), pp. 874–901.
- [198] Hongyang Li et al. “Is cobalt needed in Ni-rich positive electrode materials for lithium ion batteries?” In: *Journal of The Electrochemical Society* 166.4 (2019), A429.
- [199] Alex Zunger et al. “Special quasirandom structures”. In: *Physical Review Letters* 65.3 (1990), p. 353.