

Simulation of Defects in Molecular Materials using ab initio Methods

by

Eoghan Gormley

A dissertation accepted and approved in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy
in Chemistry

Dissertation Committee:

Carl Brozek, Chair

Christopher Hendon, Advisor

Victoria DeRose, Core Member

Benjamin McMorran, Institutional Representative

University of Oregon

Summer 2024

© 2024 Eoghan Gormley
This work is openly licensed via CC BY-NC-ND 4.0.

DISSERTATION ABSTRACT

Eoghan Gormley

Doctor of Philosophy in Chemistry

Title: Simulation of Defects in Molecular Materials using *ab initio* Methods

Molecular materials such as metal-organic frameworks (MOFs) are emerging as an interesting class of materials for their wide range of applications as well as their customizability and tunability. Like other materials, the properties of molecular materials can be modified through the use of defects - however, the nature of these defects in these systems can be much more complex than in conventional materials. Because of this, modifying conventional computational methods is essential for understanding defects in these systems. This research addresses the limitations of traditional approaches by refining techniques for calculating chemical potentials and employing higher levels of theory, such as many-body perturbation theory (MBPT), to model the unique electronic structures of MOFs. These modifications attempt to account for the intricate chemical environments, large unit cells, and significant structural dynamics inherent to molecular materials. We propose that chemical potentials for defects should be calculated based on the reactions that form the defects, thereby incorporating the thermodynamics of covalent bond formation and cleavage, as opposed to referencing them to elemental phases and treating all atoms of a given species as thermodynamically equivalent. We also attempt to use MBPT to more accurately predict the electronic properties of these materials (thereby overcoming the underestimations of fundamental band gaps typical of simpler *ab initio* methods) which is essential to predict the stability of defects that these materials

can host. Finally, a case study on Yttrium-based MOFs with HHTP linkers demonstrated the potential impact of defects and guest ions on conductivity, showcasing the practical implications of accurate defect modeling. Theoretical findings enhance the predictive power of computational methods for MOFs, informing material design and optimization for applications in catalysis, gas storage, and electronic devices. This dissertation includes previously published coauthored material.

CURRICULUM VITAE

NAME OF AUTHOR: Eoghan Gormley

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR, USA
Western Washington University, Bellingham, WA, USA

DEGREES AWARDED:

Master of Science, Chemistry, 2021, University of Oregon
Bachelor of Music, Vocal Performance, 2018, Western Washington University
Bachelor of Science, Mathematics, 2018, Western Washington University

AREAS OF SPECIAL INTEREST:

Computational Chemistry
Quantum Chemistry
Physical Chemistry

PROFESSIONAL EXPERIENCE:

Research Assistant and Teaching Assistant, University of Oregon, Eugene,
OR, USA
Front-End Supervisor, Ross Dress for Less, Vancouver, WA, USA
Math Fellow, Western Washington University, Bellingham, WA, USA

GRANTS, AWARDS AND HONORS:

NSF Graduate Research Fellowship Position (Honorable Mention), 2021
Western Washington University Dean's List, 2013-2018
College Board National AP Scholar, 2013
WMEA State Solo & Ensemble Competition - 2nd place for Tenor Voice

PUBLICATIONS:

- Gormley, E. L.**, Berger, R. F., Binding maps for the study and prediction of bimetallic catalyst surface reactions: The case of methanol oxidation, *J. Inter. Quant. Chem.*, **2018**
- Fabrizio, K., **Gormley, E. L.**, Davenport, A. M., Hendon, C. H., Brozek, C. K., Gram-scale synthesis of MIL-125 nanoparticles and their solution processability, *Chem. Sci.*, **2023**
- Kadota, K., Chen, T., **Gormley, E. L.**, Hendon, C. H., Dincă, M., Brozek, C. K., Electrically conductive [Fe₄S₄]-based organometallic polymers, *Chem. Sci.*, **2023**
- Jiang, Y., Lee, S. B., Herzog-Arbeitman, J., Yu, J., Feng, X., Hu, H., Călugăru, D., Brodale, P. B., **Gormley E. L.**, Vergniory M. G., Felser, C., Blanco-Canosa, S., Hendon, C. H., Schoop, L. M., Bernevig, B. A., Pb₉Cu(PO₄)₆(OH)₂: Phonon bands, localized flat-band magnetism, models, and chemical analysis, *Phys. Rev. B*, **2023**
- Welty, C., **Gormley, E. L.**, Oppenheim, J., Dinca, M., Hendon, C. H., Stadie, N., On the Divergent Electrically Conductive Pathways in Yttrium-Based 2- and 3-Dimensional Metal-Organic Frameworks, *ACS Materials Lett.*, **2024**

ACKNOWLEDGEMENTS

First and foremost, I would like to express my deepest gratitude to my advisor, Professor Christopher Hendon. Thank you for taking me into your research group at a critical time in my academic journey and for providing a space where I could learn and accomplish a great deal in a short period. Your guidance, patience, and support have been invaluable to me throughout this process. I would also like to extend my sincere thanks to the members of my dissertation committee: Professors Carl Brozek, Vickie DeRose, and Ben McMorran. Your insightful questions and unwavering support have been instrumental in shaping my research and guiding my thinking. To Professor Mark Lonergan, I am deeply appreciative of your willingness to listen during the many times I found myself overwhelmed, and for the sage advice you offered when I needed it most. Your kindness and understanding have meant more to me than words can express.

I owe a great deal of thanks to my incredible healthcare team for helping me navigate the myriad of health challenges I faced during my time in grad school. It's remarkable how many chronic health conditions can come to light under the stress of a PhD, and your expertise and care have been crucial to my success. I would also like to acknowledge the wonderful therapists I have worked with during my graduate studies, Alyssa Martin and Jennifer Beltz. The journey to a PhD can be particularly challenging for mental health, and your support in helping me process and manage my feelings has been invaluable.

To my parents, Sean and Sheri Gormley, thank you for your unwavering financial and emotional support. Dad, thank you for always being available to console me when I'm on the verge of a panic attack or other form of meltdown.

Lastly, to all the amazing friends I made during my time at the University of Oregon, thank you for the love and support you have given me. Your presence in my life has made this journey not just bearable but also enriching.

TABLE OF CONTENTS

Chapter	Page
I. INTRODUCTION	17
1.1. Defects in materials	17
1.2. Defining the thermodynamic nature of the defect of charge q	18
1.3. The chemical potential: the source of the defect	19
1.3.1. The reservoir and phase stability.	22
1.3.2. Choosing a reference for the chemical potential	24
1.3.3. Ternary phase diagrams for defect stability	26
1.4. The Fermi level: the chemical potential for electrons	27
1.4.1. Charge States and Thermodynamic Transition Levels	28
1.4.2. The importance of accurate band edges	31
1.4.3. Shallow Levels and Deep Levels	33
1.5. Correcting for self-interaction in periodic calculations	34
II. METHODOLOGY	38
2.1. Ground State Electronic Structure Methods	38
2.1.1. The many-body problem	38
2.1.2. Density-functional theory as an alternative to the many-body Schrödinger equation	39
2.1.3. The Hohenberg-Kohn theorems enable approximations allowing the many-body problem to be solved	40
2.2. Types of Functionals used in DFT Calculations	41
2.2.1. Pure functionals	41
2.2.2. Hybrid functionals	43

Chapter	Page
2.3. Implementing DFT using VASP and Quantum Espresso	44
2.3.1. Plane-Wave DFT	44
2.3.2. VASP: The State of the Art Periodic DFT Software	46
2.3.3. Quantum Espresso: An Open Source DFT Package	47
2.4. Excited State Electronic Structure Methods	49
2.4.1. Many-Body Perturbation Theory	49
2.4.2. The <i>GW</i> Approximation	51
2.4.3. The Bethe-Salpeter Equation	54
III. CHEMICAL POTENTIALS	57
3.1. Overview: Chemical Potentials for Molecular Materials	57
3.2. Crystalline tetrahydrofuran with interstitial tetrahydrothiophene: an O _S substitution?	59
3.3. Defects from Molecular Sources	61
3.3.1. Hydrogen from a molecular source	61
3.3.2. Hydrogen Defects in a Titanium-Based MOF	66
3.4. Discussion	70
IV. BAND GAPS	71
4.1. Current Methods Underestimate Fundamental Band Gaps of MOFs	71
4.2. Motivation: Previous Studies on Organic-Inorganic Halide Perovskites	73
4.3. Many-Body Perturbation Theory as a way to accurately obtain E_g and E_{opt}	76
4.3.1. Challenges of applying MBPT to MOFs	76
4.3.2. Methods to improve computational efficiency	78
4.4. Results and Discussion	81
4.5. Future Directions	84

Chapter	Page
V. APPLICATION: ELECTRONIC DEFECTS MEDIATED THROUGH GUEST IONS	88
5.1. Divergent electrically conductive pathways in 2- and 3-dimensional yttrium-based MOFs	88
5.2. Methods Section	96
5.2.1. Synthesis of YHOTP and Y ₆ HOTP ₂	96
5.2.2. Electrical Conductivity Measurements	97
5.2.3. Materials Characterization	97
5.2.4. Computational Methods	97
VI. CONCLUSION	99
APPENDIX: SUPPLEMENTARY INFORMATION FOR CHAPTER V	104
REFERENCES CITED	110

LIST OF FIGURES

Figure	Page
<p>1. (A) The standard technique used to calculate chemical potentials applied to CuPb substitutions in lead hydroxyapatite. Cu-rich (lead-rich) conditions correspond the energy of a single atom from a bulk Cu (Pb) unit cell, while poor conditions were referenced to the energy of CuO (PbO) by subtracting the energies of the oxygen atoms (which are each assigned half the energy of an O₂ molecule). (B) The same technique applied to interstitial hydrogen in FAPbI₃, where μ_{H} is referenced to twelve possible source molecules, and the non-hydrogen atoms in the source molecules are assigned the energy of a single atom in their corresponding elemental phases.</p>	20
<p>2. A ternary phase diagram giving the relative chemical potential of hydrogen as a function of the relative chemical potentials of zinc and oxygen.</p>	27
<p>3. Formation energy plots of the types of interstitial hydrogen defects observed in ZnO. This plot indicates that H_i⁺ defects are the most stable defect (since they have the lowest energy). The charge transition above the CBM can be attributed to hydrogen-induced resonance.</p>	29

Figure	Page
<p>4. Adapted from Fig. 2 in Reference 18. (a) Monopole corrections were a simple starting point for charged defect corrections, as they can model a wide range of systems. (b) The multipole correction was introduced to account for more strongly correlated systems whose interactions lead to notable structural distortions. (c) The FNV scheme, which models the defect charge density as a gaussian, can be applied to a wide range of systems such as Eu_{Zn} defects in ZnS. (d) The scheme proposed by Kumagai and Oba uses atomic site potentials to screen the gaussian charge density anisotropically and calculates the alignment from an average rather than aligning the potential to the periodic boundary. (e) Numerically calculating the correction term from the self-consistent charge density is widely applicable, and inherently contains all information about the jellium charge and screening without needed additional approximations. (f) The scheme of Da Silva et al. makes the assumption of isotropic screening, but can be directly and efficiently implemented in widely-used DFT</p>	36
<p>5. Schematic of a THT molecule replacing a THF molecule in a THF molecular crystal</p>	60
<p>6. A single imaginary frequency at the Γ-point corresponding to an unoptimized optical phonon mode in crystalized THF with 50% substitution of THT. Such an optical mode did not show up in pristine THF, pristine THT, or in THT with 50% substitution of THF, indicating that 50% substitution of THT in the THF lattice is not dynamically stable.</p>	62
<p>7. Predicted number of H_i defects in MIL-125 as a function of the (relative) chemical potential of hydrogen. The dashed lines depict chemical potentials calculated from various alcohols using Eqs. 3.13 (left) and 3.14 (right).</p>	67
<p>8. Predicted number of H_i defects in MIL-125 as a function of the (relative) chemical potential of hydrogen. The chemical potentials are calculated from 9,10-DHA and 1-butene using the two schemes discussed in the text.</p>	69

Figure	Page
9. Simulated absorption curves for MOF-5, MIL-125, and UiO-66 calculated using the BSE. The optical band gaps (dashed lines) are determined by the lowest-energy exciton from the calculation. The gaps for MOF-5 and UiO-66 are lower than the experimental values, likely because they are underconverged with respect to the self-energy. However, the gap for MIL-125 exceeds the experimental optical gap, likely due to significant spin-orbit coupling near the conduction band minimum.	83
10. Materials characterization comparison of YHOTP and Y ₆ HOTP ₂ . (a) Powder XRD patterns reveals the microcrystalline sample is highly crystalline. (b) Pressed pellet conductivities of five different batches of each MOF. (c) XPS show the presence of Cl in the Y ₆ HOTP ₂ , and differences in the C 1s binding energy, attributed to differences in linker oxidation state.	91
11. (a) The hexagonal (YHOTP) and (b) cubic (Y ₆ HOTP ₂) crystal structures and corresponding electronic band structures. The former has pathways of through-space π -interactions, while the latter features isolated chemical motifs. The proximity of π -clouds in YHOTP is reflected in its electronic band structure, where it is predicted to be an out-of-plane metal and in-plane narrow gap insulator. Y ₆ HOTP ₂ is predicted to be a narrow gap insulator due to limited spatial overlap throughout the crystal, with band edge densities localized on the ligand. Charge carrier effective masses are labeled. C, O, Cl, H, and Y, are depicted in black, red, green, white, and sage, respectively.	93
12. A comparison of the hexagonal (YHOTP) and cubic (Y ₆ HOTP ₂) ligand charge transfer interactions. The electron coupling (ET) is consistently larger than the hole coupling (HT) driven by the propensity for the ligands to rearomatize. C, O, and H are depicted in black, red, and white, respectively.	94
A.13. EPR spectra of YHOTP and Y ₆ HOTP ₂ between 325-350 mT, normalized by mass.	106
A.14. Iterative two-point probe measurements of packed pellet resistance of YHOTP (I-V curve swept from -1.0 to 1.0 V), upon hand tightening between measurements.	107

Figure	Page
A.15. Molecular orbitals of HOTP placed at positions consistent with Y6HOTP2, consisting of symmetric (bottom) and antisymmetric (top) combinations of the individual linker molecular orbitals. The antisymmetric state is higher in energy, suggesting J-like aggregation behavior.	108
A.16. A comparison of the hexagonal (YHOTP) and cubic (Y6HOTP2) ligand charge transfer interactions calculated using the ZINDO method. The electron and hole couplings were calculated using the open-source code py-MOO. The HT values are comparable to Figure 3, but this method calculated smaller ET values. Despite this, the ET values are still consistently larger than the HT values.	109

LIST OF TABLES

Table	Page
1. Comparison of the <i>ab initio</i> (absolute) hydrogen chemical potentials referenced to select alcohols using Eqs. 3.13 and 3.14. All values are given in units of eV.	68
2. Calculated band gaps for MOF-5, MIL-125, and UiO-66. All values are given in eV.	82
A.3. Structural models for YHOTP and Y ₆ HOTP ₂ . Coordinated water molecules (identified as one-sided O atoms in single-crystal studies) are included. Charge balance is indicated based on EPR results.	104
A.4. Single-crystal XRD structure refinement of Y ₆ HOTP ₂	105
A.5. Chemical composition of YHOTP and Y ₆ HOTP ₂ (in at%, excluding H) determined by XPS analysis.	105
A.6. Chemical composition of YHOTP and Y ₆ HOTP ₂ (in at%, excluding H) determined by combustion analysis (C, N), inductively-coupled plasma (ICP) analysis (Y, Cl), and the difference (O).	105
A.7. Model chemical composition of YHOTP and Y ₆ HOTP ₂ (in at%, excluding H) based on the combined experimental analyses: XPS, CHN, ICP-OES, and XRD.	106
A.8. Chemical composition of YCl ₃ ·6H ₂ O (in at%) determined by energy dispersive X-ray spectroscopy (EDX) using a 10 kV accelerating voltage.	106

CHAPTER I

INTRODUCTION

Figures 1 and 4 in this chapter are taken or modified from unpublished work that was co-authored by Christopher H. Hendon.

1.1 Defects in materials

Defects in materials play a critical role in determining the properties and functionalities of various materials used in modern technology. For instance, doping silicon with impurities is a fundamental process in the semiconductor industry, enabling the production of p-type and n-type semiconductors, which are essential for the functionality of electronic devices like transistors and diodes.¹ The study of defects is not limited to semiconductors but extends to a wide range of materials, including metals, ceramics, and two-dimensional materials, where defects can significantly affect mechanical properties, electrical conductivity, and thermal stability. Understanding and modeling these defects at the atomic level is crucial for the design and optimization of materials with desired properties.²

Density Functional Theory (DFT) has become a powerful tool for investigating the properties of defects in materials due to its ability to provide detailed electronic structure information. One of the key quantities of interest in defect studies is the formation energy, which can be readily calculated using DFT. The formation energy of a defect D $E_f(D^q)$ in the charge state q is given by

$$E_f(D^q) = E_{\text{Total}}^{\text{DFT}}(D^q) - E_{\text{Total}}^{\text{DFT}}(\text{Bulk}) + \sum_i n_i \mu_i + qE_{\text{F}} + E_{\text{corr}}, \quad (1.1)$$

where $E_{\text{Total}}^{\text{DFT}}(D^q)$ and $E_{\text{Total}}^{\text{DFT}}(\text{Bulk})$ are the DFT total energies of the defective and pristine systems, respectively. The term $\sum_i n_i \mu_i$ accounts for the chemical potentials of the atoms added or removed to form the defect, qE_{F} represents the contribution of the defect charge state to the formation energy, and E_{corr} is the

so-called finite-size corrections that account for self-interaction energies of defects under periodic boundary conditions.

The ability of DFT to provide insights into the local electronic environment around defects has led to significant advances in materials science.^{1,3} Computational studies using DFT have facilitated the exploration of defect dynamics, including migration pathways and activation energies, which are critical for applications such as diffusion in alloys and ionic conductivity in solid electrolytes.^{2,4} These studies contribute to a deeper understanding of material behavior under various conditions, from everyday use to extreme environments. The ongoing development of computational techniques and increasing computational power continue to expand the capabilities of DFT, making it an indispensable tool in the study of defects in materials science.^{4,5}

1.2 Defining the thermodynamic nature of the defect of charge q

The first two terms in Eq. 1.1 are fairly self-explanatory. The difference between these two terms constitutes the so-called raw defect formation energy, which is simply the change in the total energy of the material due to the presence of the defect. There are three calculations necessary to compute the raw defect formation energy:^{6,7}

1. Fully optimize the primitive unit cell of the pristine host material, allowing both the volume of the unit cell and the ion locations to relax to their lowest-energy configuration.
2. Fully optimize a supercell of the pristine host material. The lattice vectors of this supercell should be as close to the same length as possible to minimize anisotropic interactions between neighboring cells, and each vector should be at least 10 \AA long.

3. Using the optimized supercell from the previous step, introduce the desired defect(s) to this cell and optimize under constant volume and with fixed lattice vector lengths.

The raw defect formation energy is the difference in energy calculated from the last two steps. For all calculations charge and spin states should be carefully considered, as the defect may introduce additional unpaired electrons to the system.

1.3 The chemical potential: the source of the defect

Chemical potential plays a key role in influencing the movement of atoms or molecules within a system, acting as the driving force behind this process. In the realm of materials, it signifies the amount of energy required to either add or remove an atom or molecule from a chemical system. When defects arise in a material, such as vacancies (missing atoms) or interstitials (extra atoms), the formation energy pertains to the energy needed to create these defects. The chemical potential considers the energy necessary to eliminate these defect species from a reactant (or in the case of vacancies, the energy gained by these defect species transitioning to another product). Chapter III will discuss modifying conventional methods for determining chemical potentials to better suit molecular materials, and is based on unpublished work that was co-authored by Christopher H. Hendon.

Typically, chemical potentials are referenced to elemental substances or bulk materials^{1,8,9} For instance, when examining defective metal atoms, the chemical potential will be calculated from the energy of the pure metal in its most stable state. This is often designated as the “rich” condition for that species. In scenarios where the defect is more likely to come from a compound (such as Cu from CuO), the convention is to say that species in a “poor” condition (Figure 1a).

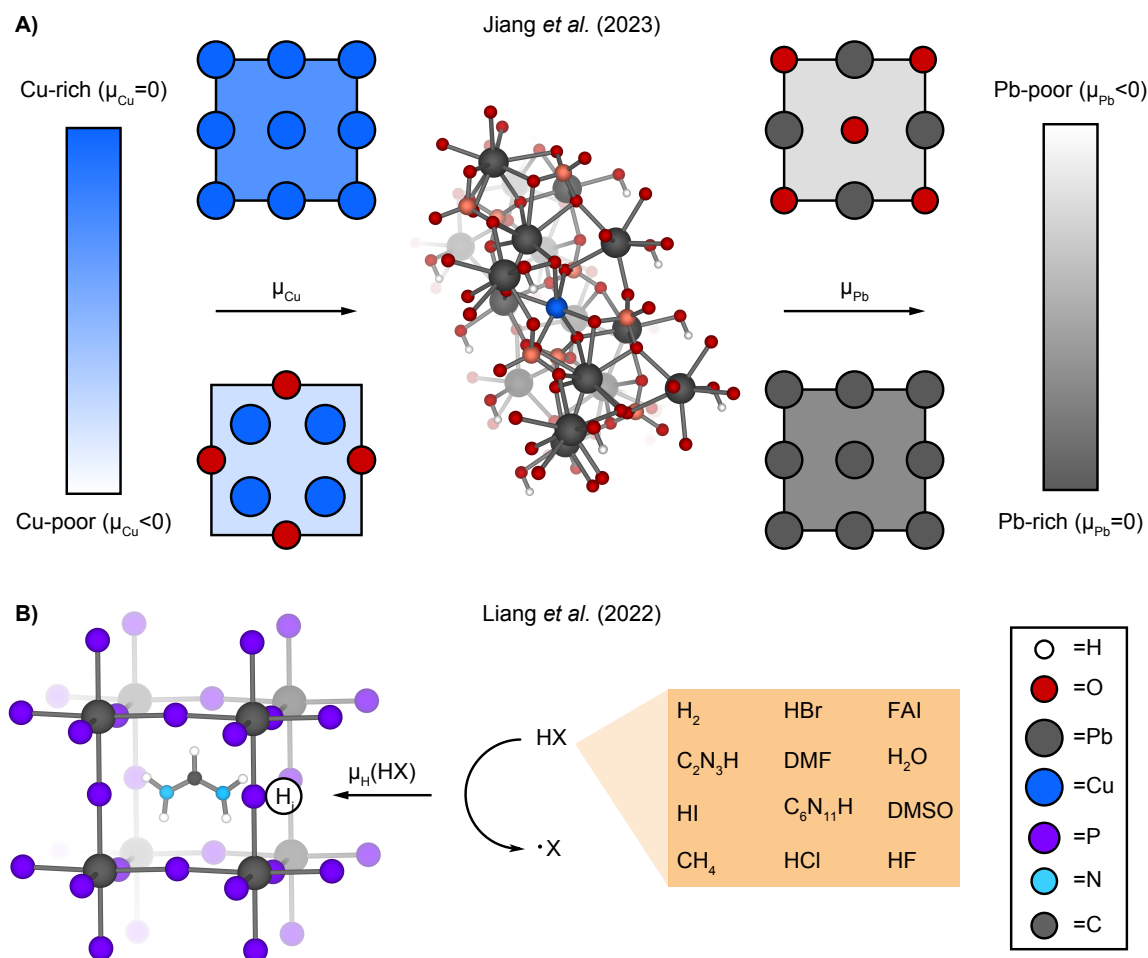


Figure 1. (A) The standard technique used to calculate chemical potentials applied to CuPb substitutions in lead hydroxyapatite. Cu-rich (lead-rich) conditions correspond to the energy of a single atom from a bulk Cu (Pb) unit cell, while poor conditions were referenced to the energy of CuO (PbO) by subtracting the energies of the oxygen atoms (which are each assigned half the energy of an O₂ molecule). (B) The same technique applied to interstitial hydrogen in FAPbI₃, where μ_{H} is referenced to twelve possible source molecules, and the non-hydrogen atoms in the source molecules are assigned the energy of a single atom in their corresponding elemental phases.

Contrary to many traditional semiconductors, molecular materials exhibit intricate and often complex structures that are both large and flexible. The chemical environment within these materials can vary significantly across different regions, affecting the chemical potential and complicating the determination of accurate values for specific areas. Interstitial defects in molecular materials can often come from molecular sources, such as solvent molecules like DMF or ethanol, thus necessitating the destruction and formation of covalent bonds.¹⁰ The typical procedure for calculating chemical potentials referenced to molecular sources is to decompose the molecule into its constituent atoms, and then reference the energies of the non-defect atoms to their elemental phases. For example, Liang et al. calculated the chemical potential μ_{H} for interstitial hydrogen atoms in formamidinium lead triiodide (FAPbI₃) referenced to twelve molecular sources (Figure 1b).⁸ This approach inherently assumes that all hydrogens in the molecule are chemically equivalent, which may be true for simple molecules like methane but for more complex molecules like DMF this may not be sufficient. Additionally, this approach fails to account for the stability of the product molecules formed after the hydrogen is lost – removing a hydrogen atom from methane would not yield a stable product.

Chemical potentials are essential for understanding defect formation in molecular materials, but computing them is challenging and requires care. It is important to understand not only the chemical environment of the bulk system, but also the environments of the defect source molecules and the chemical reactions (i.e. reactants and products) involved in the exchange of defects with the bulk system. By addressing these hurdles head on, there is potential to fully exploit the

capabilities of molecular materials across various applications, from catalysis to gas storage and beyond.

1.3.1 The reservoir and phase stability. The chemical potential is a powerful tool for predicting defect stability, especially in a computational context. To understand the idea behind this quantity, imagine a mixture of all the species (atoms, ions, etc.) involved in synthesizing the desired product floating around freely in a container of some sort, and that this mixture has a total energy associated with it. In order to form a product, one must remove the appropriate species from this container, and as a result the total energy of the mixture will decrease (we will ignore the interactions between the species in the mixture for simplicity). We call this hypothetical mixture that contains all the species needed to synthesize a desired product the reservoir. It should be emphasized that the reservoir is a hypothetical space where the reactants decompose completely into their constituent atoms, and therefore it does not necessarily contain the reactants for a chemical reaction. This concept provides a simple and intuitive way to define the so-called chemical potential for a given species - it is the energy lost from the reservoir when one particle of that species is removed. In other words, it represents the energy contribution of that particle to the total energy of the reservoir. Thermodynamically, this can be written in terms of the internal energy of the reservoir U_{res} :

$$\mu_i = \left(\frac{\partial U_{\text{res}}}{\partial n_i} \right)_{S,V,n_j \neq n_i}, \quad (1.2)$$

where μ_i is the chemical potential of species i , and n_i is the number of particles of species i being removed. Notice that if the total energy of the reservoir is a positive number, this definition requires that the chemical potential be a negative number.

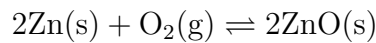
In practice, it is more useful to define the chemical potential in terms of the change

in enthalpy or the Gibbs free energy associated with removing the species from the reservoir using the appropriate Legendre transformations:

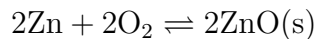
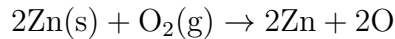
$$\mu_i = \left(\frac{\partial H_{\text{res}}}{\partial n_i} \right)_{S,P,n_j \neq n_i} = \left(\frac{\partial G_{\text{res}}}{\partial n_i} \right)_{T,P,n_j \neq n_i} . \quad (1.3)$$

This quantity allows one to explore different phases of the species in the reservoir under given thermodynamic conditions. In this case, a phase refers to a stable form that species in the reservoir can leave as. The stability of these forms depends on the thermodynamic conditions of the reservoir's surrounding environment, and this stability is determined by the chemical potential.

As a pedagogical exercise, consider the bulk ZnO in equilibrium with bulk Zn and O₂ gas:



The reservoir in this case is a hypothetical mixture of free Zn atoms and free O atoms. We can think about this as a hypothetical intermediate reaction where the elemental Zn and O spontaneously decompose:



The different forms that the atoms can leave the reservoir as (i.e. bulk Zn, O₂, or bulk ZnO) are often referred to as phases. Based on our definition of the chemical potential, Eq. 1.2 implies that it takes less energy for the elemental phases of Zn and O to decompose into the reservoir than it does for Zn and O to leave the reservoir in their elemental phases. This yields the following conditions for the

chemical potentials of Zn and O:

$$\mu_{\text{Zn}} \leq \mu_{\text{Zn}[\text{bulk}]} \quad (1.4)$$

$$\mu_{\text{O}} \leq \frac{1}{2}\mu_{\text{O}_2} \equiv \mu_{\text{O}[\text{O}_2]} \quad (1.5)$$

where the equality conditions in Eqs. 1.4 and 1.5 signify equilibrium with the elemental phases. Since the Zn and O in the reservoir are in equilibrium with the ZnO phase, our definition of the chemical potential tells us that it takes as much energy to remove ZnO from the reservoir as it does for ZnO to decompose into the reservoir. Therefore,

$$\mu_{\text{Zn}} + \mu_{\text{O}} = \mu_{\text{ZnO}[\text{bulk}]} \quad (1.6)$$

Note that from Eq. 1.3 the change in Gibbs's free energy for this reaction is given by

$$\Delta_f G^\circ = \mu_{\text{ZnO}[\text{bulk}]} - (\mu_{\text{Zn}[\text{bulk}]} + \mu_{\text{O}[\text{O}_2]}), \quad (1.7)$$

and therefore in order for ZnO to form spontaneously we must have that

$$\mu_{\text{ZnO}[\text{bulk}]} \leq \mu_{\text{Zn}[\text{bulk}]} + \mu_{\text{O}[\text{O}_2]}. \quad (1.8)$$

In other words, in order for a phase to be stable it must be lower in energy than its constituent phases together. Note that since entropic effects are often ignored for crystalline materials, Eq. 1.7 may be expressed equivalently in terms of $\Delta_f H^\circ$.

1.3.2 Choosing a reference for the chemical potential. As with most thermodynamic quantities, the chemical potential must be defined in reference to a particular value. Typically, the chemical potential for a substance is defined using the stability of its purest form as a reference point as we have done in Eqs. (?) and (?). In other words, the chemical potential for a substance is defined relative to the energy required to remove a single atom from its elemental phase.

It should be noted here that in defect chemistry synthetic conditions are typically kept close to equilibrium so that defect concentration can be more easily controlled. For this reason, the convention is to define so-called “rich” and “poor” conditions based on the equilibrium conditions in the reservoir. We say that the system is in rich conditions for a species if the species in the reservoir is in equilibrium with its elemental phase; if the stability of the species is entirely dependent on the stability of the other phases in the reservoir we say that the system is in poor conditions for that species.

To illustrate this point, consider the example with ZnO above. If the synthetic conditions are Zn-rich, Eq. 1.4 becomes an equality since the Zn in the reservoir is in equilibrium with the bulk Zn phase. However, this implies that Eq. 1.8 must be a strict inequality in order for ZnO to form spontaneously. Therefore, in Zn-rich conditions the chemical potential for O is given by

$$\mu_{\text{O}} = \mu_{\text{ZnO}[\text{bulk}]} - \mu_{\text{Zn}[\text{bulk}]}.$$

This indicates that μ_{O} is dependent on the chemical potentials for the Zn and ZnO, which means that the system is in O-poor conditions. Similarly, if we were to choose O-rich conditions we would find that the system has to be in Zn-poor conditions.

In cases where more than two species are involved in product synthesis these definitions become ambiguous, and one must be careful to define a reference for this additional species. However in the case of binary mixtures such as ZnO, the stability of the elemental phases can give valuable insight into the types of native defects that can form as the thermodynamic stability of Zn and O can alter the stoichiometric ratio of Zn to O in the product crystal. For example, under Zn-rich conditions we would anticipate a higher concentration of Zn interstitial and antisite

defects, and a greater concentration of O vacancies due to O-poor conditions.

This begins to illustrate how altering the chemical potential via thermodynamic conditions can be a powerful tool for selective defect formation.

1.3.3 Ternary phase diagrams for defect stability. When considering non-native defects (i.e. defects of species that are not present in the pristine compound), it is important to consider the stability of all phases containing the species in the desired product. For example, when considering hydrogen defects in ZnO one should also consider the ternary phase Zn(OH)_2 . This is important because this can determine, for example, whether the defects should be thought of as interstitial hydrogen defects in a ZnO crystal or hydrogen vacancies in a Zn(OH)_2 crystal.

Similar to our treatment of ZnO, we define the boundary conditions for the stability of Zn(OH)_2 based on the equilibrium conditions for the possible phases that can be formed from Zn, O, and H. Thus, along with Eq. 1.6 we have equilibrium with bulk Zn (Zn-rich), O_2 gas (O-rich), ZnO, water, and Zn(OH)_2 :

$$2\mu_{\text{H}} + \mu_{\text{O}} = \mu_{\text{H}_2\text{O}} \quad (1.9)$$

$$\mu_{\text{Zn}} + 2\mu_{\text{O}} + 2\mu_{\text{H}} = \mu_{\text{Zn(OH)}_2} \quad (1.10)$$

$$2\mu_{\text{H}} + 2\mu_{\text{O}} = \mu_{\text{H}_2\text{O}_2}. \quad (1.11)$$

We can visualize these conditions as lines on a two-dimensional plot of the chemical potentials of Zn and O (see Figure 2). The blue line in Figure 2 is just Eq. 1.6, where μ_{ZnO} is the enthalpy of formation for ZnO. The red line is the Eq. 1.10 under H-rich conditions, meaning we set $\mu_{\text{H}} = 0$ to determine the upper bound for the formation of Zn(OH)_2 with respect to the chemical potential of hydrogen. Finally, the purple line is obtained from the second equation by solving Eq. 1.10 for

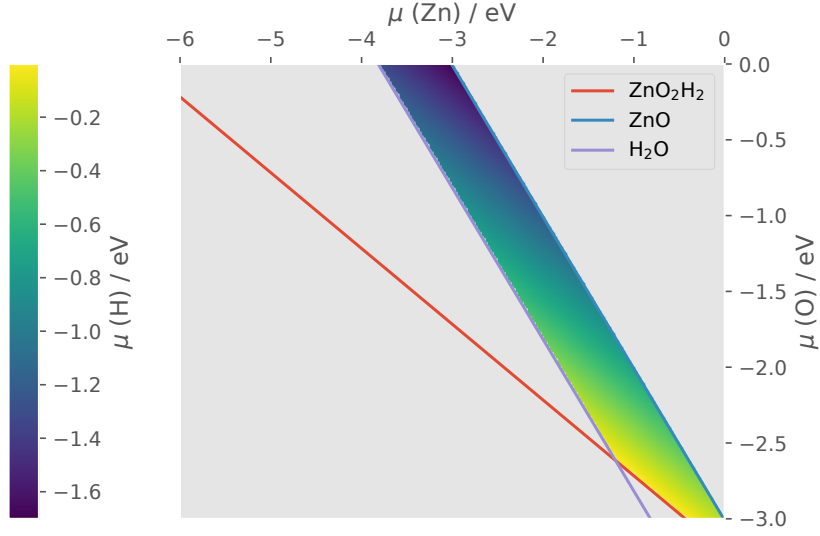


Figure 2. A ternary phase diagram giving the relative chemical potential of hydrogen as a function of the relative chemical potentials of zinc and oxygen.

μ_{H} and substituting it into Eq. 1.9, giving the equation

$$\mu_{\text{Zn}} + \mu_{\text{O}} = \mu_{\text{Zn}(\text{OH})_2} - \mu_{\text{H}_2\text{O}}. \quad (1.12)$$

Note that this equation and Eq. 1.6 have the same linear coefficients on μ_{Zn} and μ_{O} , hence why the blue and purple phase lines in Figure 2. The right side of this equation is also physically significant, as it implies that one can obtain ZnO by removing a water molecule from $\text{Zn}(\text{OH})_2$.

1.4 The Fermi level: the chemical potential for electrons

One can consider the Fermi level E_F in Eq. 1.1 to be an analogue of the chemical potential discussed in the previous section specifically for electrons. However, unlike μ_i the Fermi level is a positive value by convention because electrons are negatively charged by convention. Thus, if we think of q in analogy to n_i (with the opposite sign convention), removing an electron from the reservoir ($q = -1$) corresponds to a decrease in the internal energy of the reservoir by E_F . Reference values for are straightforward to define, as the chemical potential

landscape for electrons in a host material is defined by the band structure of that material. For semiconductors and insulators, the reference value for the Fermi level ($E_F = 0$) is defined to be the valence band maximum (VBM) of the pristine material – in other words, for a material in its electronic ground state it is the energy of the electron in the occupied state with the highest energy. A shift in the Fermi level to higher energies typically corresponds to distortions in the electron chemical potential landscape due to the presence of defects in the material or experimental application of an electric field. Conceptual understanding of the Fermi level is crucial when studying defects in materials because it provides the basis for both identifying the charge state of a defect as well as characterizing its influence on the host material.

1.4.1 Charge States and Thermodynamic Transition Levels.

Charge states are an important quantity used in classifying defects and predicting their properties. Since charged defects tend to distort local electronic structure more than their neutral counterparts, one might anticipate charged defects to have a stronger influence on optical and thermodynamic properties of a material. This is compounded by the fact that charge carriers (i.e. electrons and holes) are often exchanged between the defect and the host material, which can both alter the optical properties of the host and cause the defect to transition to a different charge state. Defects are often characterized based on these charge transitions – p-type defects exhibit transitions to lower charge states (they accept electrons from the host material, creating holes), and n-type defects exhibit transitions to higher charge states (they donate electrons to the host material). The Fermi level provides a useful tool for quantitative prediction of the energies of these charge transitions – the value of E_F at which a defect goes from one charge state to another is called a

Interstitial Hydrogen (H-rich)

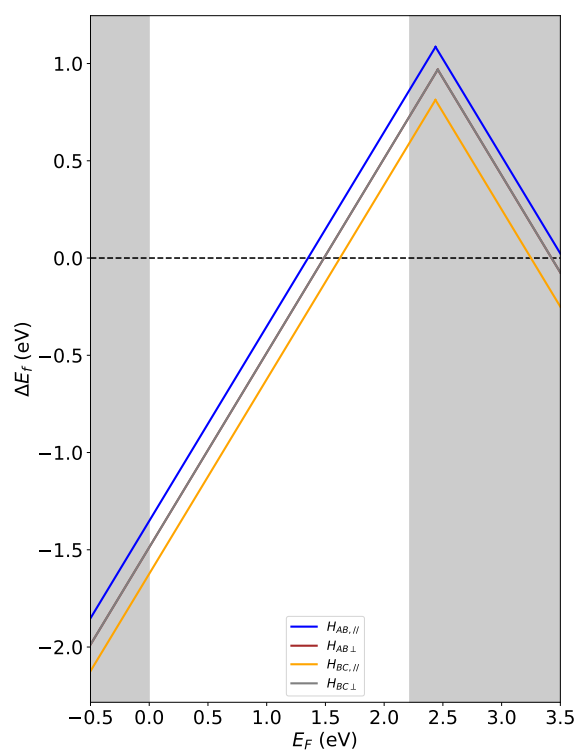


Figure 3. Formation energy plots of the types of interstitial hydrogen defects observed in ZnO. This plot indicates that H_i^+ defects are the most stable defect (since they have the lowest energy). The charge transition above the CBM can be attributed to hydrogen-induced resonance.

thermodynamic transition level. These transition levels are determined by plotting the formation energy of a defect as a function of the Fermi level (see Figure 3). In other words, if we calculate the formation energy of a defect with a charge state q with $E_F = 0$, we see that modulating the value of E_F gives the equation of a line. If we then repeat this procedure with a range of charge states, we obtain several intersecting lines. The values of E_F where these lines intersect are the thermodynamic transition levels.

To summarize this procedure mathematically, for a defect D with two different charge states q and q' , we calculate the formation energies for the defect in these charge states at the VBM to be $E^f(D^q; E_F = 0)$ and $E^f(D^{q'}; E_F = 0)$. If we plot the lines

$$E^f(E_F) = E^f(D^q; E_F = 0) + qE_F \quad (1.13)$$

$$E^f(E_F) = E^f(D^{q'}; E_F = 0) + q'E_F, \quad (1.14)$$

the value of $E^f(E_F)$ where these lines intersect can be solved for as

$$\varepsilon(q/q') = \frac{E^f(D^q; E_F = 0) - E^f(D^{q'})}{q - q'}. \quad (1.15)$$

The significance of this quantity is that it determines which charge state is most stable at a given Fermi level. If $E^f = \varepsilon(q/q')$ then charge state q is more stable, and if $E^f < \varepsilon(q/q')$ then charge state q' is more stable. For this reason, it is common to only show the most stable charge state at a given Fermi level when making plots like the one in Figure 3, which is recreated from previously published data.¹¹ The locations of these transition levels relative to both the VBM and the conduction band minimum (CBM) are used to further classify defects by their influence on the optical properties of the host material, and therefore it is

imperative to have an accurate quantitative measure of the band gap for the host material.

1.4.2 The importance of accurate band edges. For semiconductors and insulators, the band gap is typically defined to be the energetic distance between the VBM and the CBM. Physically, this corresponds to the energy needed to excite a valence electron bound to one atom in the material into a conduction band, where it is then free to move throughout the solid as a charge carrier. It is important that we distinguish between this quantity – which is also called the fundamental band gap – and the optical band gap. The latter quantity is defined to be the minimum amount of energy needed to excite a valence electron in a semiconductor or insulator and corresponds to the threshold of a material for photon absorption.

While the fundamental band gap (denoted by E_g) and the optical gap (which we denote by E_{opt}) are identical for nearly all inorganic semiconductors, for molecular materials E_g can be more than twice the size of E_{opt} . This is because electrons in molecular materials often bind very strongly to their host molecules, and therefore optical excitation of these electrons can create a bound excited state that will decay over time. This creates another important issue in the study of defects - if a defect has a unique optical signature, then one should be aware of the proximity of the defect level relative to E_{opt} (particularly in an experimental setting). The difference between the fundamental band gap and the optical band gap is called the exciton binding energy. Determination of these quantities has been a longstanding challenge in both computation and experiment.

It is well-known that the DFT method used ubiquitously in computational chemistry severely underestimates E_g due to the electron correlation functional

introducing non-physical interactions between electrons and themselves. This issue can be somewhat mitigated by using modified versions of this method, such as using extrapolation schemes or hybrid functionals.¹² Hybrid functionals have been demonstrated to accurately predict the optical band gap of a range of materials including MOFs¹³ – however, accounting for the exciton binding energy has proven to be difficult.

Currently, the best method for computationally predicting E_g is using the so-called GW approximation. This method uses quantum field theory and linear response theory to derive the propagator for an electron-hole pair using the non-interacting Green’s function (which in many-body theory refers to the same quantity as the time-correlation function). The propagator can then be used to obtain the so-called quasiparticle energies for the system, where the quasiparticles are all of the possible excitons (i.e. electron-hole pairs) for the valence states. The main appeal of this method is in the context of the G_0W_0 calculation, where the quasiparticle energies (also referred to as “self-energies”) are obtained by diagonalizing the self-energy operator Σ^{xc} in the basis of Kohn-Sham wavefunctions one can obtain through routine DFT calculations. In other words, the GW method can be used as an extension of ground-state DFT. Until recently this method has not been computationally feasible in most cases because it requires DFT calculations with many unoccupied bands. This requirement comes from the fact that the quasiparticle energies are obtained by inserting a completeness relation (i.e. a summation) into the single particle Green’s function that runs over all of the possible quasiparticle states. In principle since there are an infinite number of unoccupied states this summation would require an infinite number of terms to converge, but in practice a large number will usually suffice.

Bethe-Salpeter Equation (BSE) calculations are an extension of the *GW* method that use the quasiparticle energies from a *GW* calculation to compute the two-point linear response function for the electron-hole pairs. The two-point linear response function can then be used to account for electron-hole interactions in the quasiparticle self-energies, thus giving a more reliable value for E_{opt} . Thus, the ideal workflow for studying the Fermi energy term in Eq. 1.1 would be to first perform a routine DFT optimization on the pristine material, then perform a *GW* calculation to predict E_g using the results of that calculation, and then finally run a BSE calculation to predict the optical signature E_g of the pristine material.

1.4.3 Shallow Levels and Deep Levels. We now arrive at one of the most important parts of our discussion on the Fermi level: classifying thermodynamic transition levels. This is arguably the most important step in determining whether defects show promise for use in devices and other applications because the charge transitions introduced by a defect correspond to thermal ionization energies. In other words, the location of these transition levels relative to the band edges determines how likely it is for a defect to be thermally ionized at room temperature, and we again emphasize that quantitatively accurate prediction of the band edges is essential for defect classification. If a transition level is close to the band edges (either the VBM or CBM), it is called a shallow level. Likewise, if a transition level is far from both of the band edges (so closer to the center of the band gap), it is called a deep level.

Shallow levels are used to determine n-type versus p-type conductivity (excess of holes versus excess of electrons, respectively). The proximity of shallow levels to the band edges means that the defect is very likely to be thermally ionized at room temperature because the energetic distance is small compared to $k_{\text{B}}T$. If

the transition level is close to the conduction band, ionization of electrons from the valence band to the defect will occur more efficiently and therefore leave holes in the valence band (p-type). Similarly, if the transition level is close to the CBM then electrons at the defect level will ionize more efficiently into the conduction band, thus creating mobile charge carriers more easily (n-type). Because of this, defects with shallow levels are much more favorable in applications to electronic devices. On the other hand, in deep levels the energy to mobilize charge carriers is much larger than $k_B T$, and therefore it is less likely for the defect to be thermally ionized at room temperature. This results in semi-insulating behavior in the material, which is typically undesirable for applications involving conductivity but desirable for their insulating properties. Chapter V is based on work analyzing the affect of defect-introduced charge carriers in a 3-dimensional flat-band MOF. This work was co-authored with Connor Welty, Julius J. Oppenheim, Mircea Dincă, Christopher H. Hendon, and Nicholas P. Stadie.

1.5 Correcting for self-interaction in periodic calculations

One of the most important and ubiquitous applications of defects is doping materials to mobilize charge carriers in the conduction and/or valence bands^{14,15}. As discussed in the previous section, defects introduce midgap states that can either donate or accept electrons to or from the host material respectively. In either case, this exchange of electrons leaves the defect with a net charge. In nature, these types of defects are usually charge balanced by counter ions (forming what is known as a Frenkel pair), but these counter ions can vary widely and can be separated over large distances^{16,17}. Computationalists are generally interested in studying the properties of a single, isolated defect in a periodic system, and therefore must seek corrections for the electrostatic interactions that occur between

defects in neighboring unit cells under periodic boundary conditions. Additionally, since periodic DFT calculations compute energies based on charge density (and therefore require a net zero charge on a system calculate finite energies), a fictitious “jellium” background potential is introduced that must be corrected for. This is the purpose of the final term in Eq. 1.1.

In theory, the most logical way to mitigate interactions between charged defects would be to separate them over very large instances (essentially creating an isolated defect in an infinite crystal). Obviously this approach is computationally infeasible, so the typical approach is to make a supercell of the host material and place the defect accordingly. Unfortunately, while this approach does reduce electrostatic interactions it cannot remove them entirely. A number of approaches have been proposed over the past few decades to remedy this, mainly drawing from the theories of classical electrostatics.^{18–24} These approaches have also been applied to a wide-range of systems (see Figure 4).^{25–30}

Over the last decade the most popular method of correction has been that of Freysoldt, Neugebauer, and Van de Walle (the FNV scheme).²¹ Rather than calculating the energy correction from a monopole or multipole expansion around the defect site, the FNV scheme instead models the charge density of the defect using a gaussian charge distribution centered on the defect. This method uses classical electrostatics to derive an analytical solution to the Poisson equation, and then calculates the interaction energy from the potential as the charge correction. The main challenge of this approach is that it requires calculation of the dielectric constant for the host material to use for the screening. This requires performing a phonon calculation, which can be computationally expensive for more complex

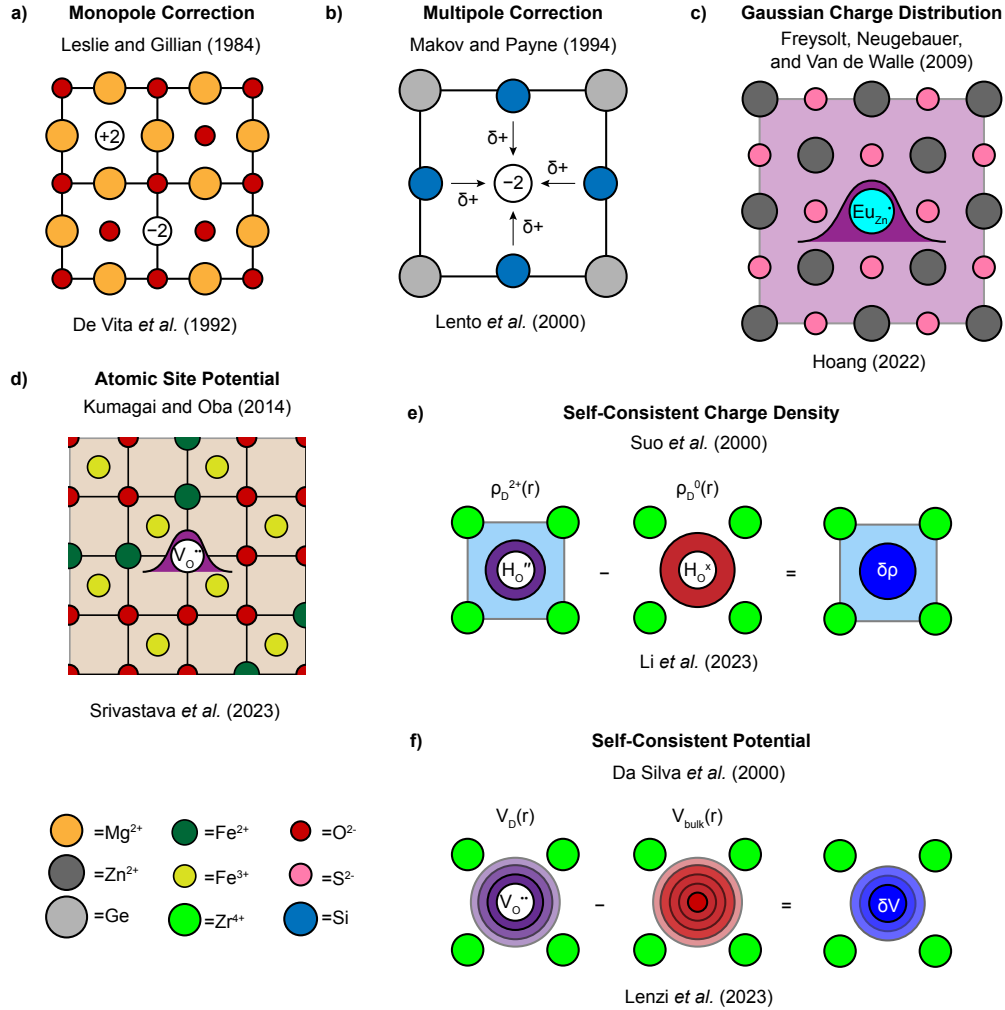


Figure 4. Adapted from Fig. 2 in Reference 18. (a) Monopole corrections were a simple starting point for charged defect corrections, as they can model a wide range of systems. (b) The multipole correction was introduced to account for more strongly correlated systems whose interactions lead to notable structural distortions. (c) The FNV scheme, which models the defect charge density as a gaussian, can be applied to a wide range of systems such as Eu_{Zn} defects in ZnS. (d) The scheme proposed by Kumagai and Oba uses atomic site potentials to screen the gaussian charge density anisotropically and calculates the alignment from an average rather than aligning the potential to the periodic boundary. (e) Numerically calculating the correction term from the self-consistent charge density is widely applicable, and inherently contains all information about the jellium charge and screening without needed additional approximations. (f) The scheme of Da Silva et al. makes the assumption of isotropic screening, but can be directly and efficiently implemented in widely-used DFT

materials. However, the main appeal of this method is its simplicity, as once the dielectric constant is obtained one may use the open access scripts available at.

More recent publications have formulated correction schemes that analyze the data from the scf calculations directly. The scheme proposed by da Silva et al computes the energy correction using the difference in charge density between the defective material and the pristine material to numerically solve the Poisson equation, and then uses this to obtain a difference in potential.²⁴ Unfortunately, this method also requires the computation of a dielectric constant. In contrast, the method developed by Suo et al takes advantage of an image-charge correction that circumvents the use of a dielectric function, thus allowing them to obtain the energy correction from only the difference in self-consistent charge density.²³

CHAPTER II
METHODOLOGY

2.1 Ground State Electronic Structure Methods

2.1.1 The many-body problem. The many-body problem lies at the heart of quantum chemistry, encapsulating the complexities and challenges of accurately describing the behavior of a system consisting of multiple interacting particles. In quantum chemistry, understanding the electronic structure of atoms, molecules, and condensed matter systems necessitates solving the Schrödinger equation for a multitude of electrons and nuclei. This task is formidable due to the exponential increase in computational complexity with the number of particles involved, leading to what is known as the “curse of dimensionality”.^{31–34} Traditional methods often fall short in providing precise solutions, prompting the development of various approximation techniques and computational strategies. These approaches, ranging from Hartree-Fock and DFT to coupled-cluster methods and quantum Monte Carlo simulations, seek to balance accuracy and computational feasibility.

The many-body problem in electronic structure theory seeks solutions to the Schrödinger equation for the N -electron Hamiltonian

$$H = \hat{T} + \hat{V} + \hat{U}, \quad (2.1)$$

where

$$T = -\frac{\hbar^2}{2m} \sum_{i=1}^N \nabla_i^2 \quad (2.2)$$

is the kinetic energy operator,

$$V = \sum_{i=1}^N V(\mathbf{r}_i) \quad (2.3)$$

is the potential energy operator due to the electrons interacting with the positively charged nuclei (where \mathbf{r}_i denotes the position operator for the i th electron), and

$$U = \sum_{i < j}^N U(\mathbf{r}_i, \mathbf{r}_j) \quad (2.4)$$

is the potential energy due to electron-electron interaction. Note that both V and U are made up of Coulomb potential energies.

It is well-known that the equation $H|\Psi\rangle = E|\Psi\rangle$ has no analytical solutions for $N > 1$ due to a multitude of reasons (such as nonlinearity and coupling of the resulting partial differential equations due to the electron-electron interactions).³⁴⁻³⁸ As a result, numerous approximations have been made to circumvent this issue. This chapter delves into the intricacies of the many-body problem and summarizes how DFT is used to approximate solutions in a way that drives progress in quantum chemistry.

2.1.2 Density-functional theory as an alternative to the many-body Schrödinger equation. Density Functional Theory (DFT) has emerged as a powerful and widely-used alternative to directly solving the many-body Schrödinger equation in quantum chemistry. Instead of attempting to solve the complex many-body wavefunction, DFT focuses on the electron density, a significantly simpler function of only three spatial coordinates, regardless of the number of electrons. The foundational theorems of DFT, established by Hohenberg and Kohn, prove that the ground-state properties of a many-electron system are uniquely determined by its electron density.³⁹ The practical implementation of DFT is realized through the Kohn-Sham formalism, which introduces a set of non-interacting reference particles that reproduce the exact electron density of the interacting system.⁴⁰ This approach transforms the original many-body problem into a more tractable form, involving a set of single-particle equations known as the

Kohn-Sham equations:

$$\left(-\frac{\hbar^2}{2m}\nabla^2 + V_{\text{ext}}(\mathbf{r}) + V_{\text{H}}[\rho](\mathbf{r}) + V_{\text{XC}}[\rho](\mathbf{r})\right)\psi_i(\mathbf{r}) = \epsilon_i\psi_i(\mathbf{r}), \quad (2.5)$$

where $V_{\text{ext}}(\mathbf{r})$ is the external potential (notated by V in Eq. 2.1), $V_{\text{H}}[\rho](\mathbf{r})$ is the Hartree potential, and $V_{\text{XC}}[\rho](\mathbf{r})$ is the exchange-correlation potential (where $V_{\text{ext}}(\mathbf{r}) + V_{\text{XC}}[\rho](\mathbf{r})$ is encompassed by U in Eq. 2.1). These equations are solved iteratively to obtain the self-consistent electron density $\rho(\mathbf{r})$, providing a balance between computational efficiency and accuracy. DFT has thus become indispensable in studying the electronic structure of atoms, molecules, and solids.⁴¹

2.1.3 The Hohenberg-Kohn theorems enable approximations allowing the many-body problem to be solved. The Hohenberg-Kohn theorems form the theoretical foundation of Density Functional Theory (DFT) and provide a pathway to approximate solutions of the many-body problem in quantum chemistry. The first Hohenberg-Kohn theorem states that the ground-state electron density uniquely determines the external potential, $V_{\text{ext}}(\mathbf{r})$, up to a constant, and thus all ground-state properties of a system. This groundbreaking theorem shifts the focus from the many-body wavefunction, which is dependent on $3N$ coordinates for N electrons, to the electron density, a function of only three spatial coordinates.³⁹ By demonstrating that the electron density contains all the information necessary to describe the ground state of a system, the first theorem significantly reduces the complexity of the problem.

The second Hohenberg-Kohn theorem introduces a variational principle for the electron density, which states that the correct ground-state density minimizes the energy functional. This functional, $E[\rho]$, is given by:

$$E[\rho] = T_s[\rho] + \int V_{\text{ext}}(\mathbf{r})\rho(\mathbf{r}) d\mathbf{r} + \frac{1}{2} \int \frac{\rho(\mathbf{r})\rho(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + E_{\text{XC}}[\rho], \quad (2.6)$$

where $T_s[\rho]$ is the kinetic energy of a system of non-interacting electrons, the second term represents the classical electrostatic interaction energy (also known as the Hartree energy), and $E_{XC}[\rho]$ is the exchange-correlation energy term that encapsulates all many-body effects. The exact form of the exchange-correlation functional, $E_{XC}[\rho]$, is unknown, but various approximations such as the Local Density Approximation (LDA) and Generalized Gradient Approximation (GGA) have been developed.⁴⁰ These approximations allow for the practical application of DFT by enabling the calculation of ground-state properties with reasonable accuracy and computational efficiency. Through these approximations, the Hohenberg-Kohn theorems provide a framework that transforms the intractable many-body Schrödinger equation into a more manageable form, facilitating the study of complex quantum systems.⁴¹

2.2 Types of Functionals used in DFT Calculations

2.2.1 Pure functionals. In Density Functional Theory (DFT), pure functionals are used to approximate the exchange-correlation energy functional, $E_{XC}[\rho]$, which plays a critical role in determining the accuracy of DFT calculations. The exchange-correlation energy encompasses the complex many-body interactions among electrons that are not captured by the simpler components of the energy functional. The development of accurate approximations for $E_{XC}[\rho]$ has been a central challenge in DFT, leading to various approaches that balance computational efficiency with accuracy.

One of the earliest and simplest approximations for the exchange-correlation functional is the Local Density Approximation (LDA). In LDA, the exchange-correlation energy at each point in space is assumed to depend only on the electron density at that point, analogous to the behavior of a homogeneous electron gas.

The LDA functional is expressed as:

$$E_{\text{XC}}^{\text{LDA}}[\rho] = \int \rho(\mathbf{r}) \epsilon_{\text{XC}}^{\text{hom}}(\rho(\mathbf{r})) d\mathbf{r}, \quad (2.7)$$

where $\epsilon_{\text{XC}}^{\text{hom}}(\rho)$ is the exchange-correlation energy density of a uniform electron gas with density ρ .⁴² LDA tends to work well for systems with slowly varying electron densities, such as bulk metals, but can be less accurate for systems with rapidly varying densities, such as molecules and surfaces.

To improve upon the LDA, the Generalized Gradient Approximation (GGA) includes not only the local electron density but also its gradient, thereby capturing more information about the inhomogeneity of the electron distribution. The GGA functional can be expressed as:

$$E_{\text{XC}}^{\text{GGA}}[\rho] = \int f(\rho(\mathbf{r}), \nabla\rho(\mathbf{r})) d\mathbf{r}, \quad (2.8)$$

where $f(\rho, \nabla\rho)$ is a function that depends on both the electron density and its gradient.⁴³ Popular GGA functionals, such as PBE (Perdew-Burke-Ernzerhof), have been designed to improve accuracy for a wide range of systems, including molecules and complex materials. GGA functionals generally provide better descriptions of binding energies, bond lengths, and other properties compared to LDA, especially for systems with significant density variations.

While both LDA and GGA are widely used and computationally efficient, they represent different levels of approximation in capturing the exchange-correlation effects. LDA's simplicity and reasonable accuracy for certain systems make it a useful tool, particularly for metallic and bulk systems. In contrast, GGA's incorporation of gradient information allows for more accurate modeling of molecular and surface phenomena, making it more versatile for a broader range of chemical and material applications.⁴⁴

2.2.2 Hybrid functionals. While pure functionals like the Local Density Approximation (LDA) and Generalized Gradient Approximation (GGA) have provided significant advancements in the practical application of Density Functional Theory (DFT), they are not without limitations. Pure functionals often struggle with accurately describing systems where electron correlation and exchange interactions play a critical role. To address these limitations, hybrid functionals have been developed, which incorporate a portion of exact exchange energy from Hartree-Fock theory with the exchange-correlation energy from DFT approximations.

Hybrid functionals improve upon the accuracy of pure functionals by mixing the exact exchange energy, E_X^{HF} , with the DFT exchange-correlation energy, $E_{\text{XC}}^{\text{DFT}}$. The general form of a hybrid functional can be written as:

$$E_{\text{XC}}^{\text{hybrid}}[\rho] = aE_X^{\text{HF}} + (1 - a)E_X^{\text{DFT}} + E_C^{\text{DFT}}, \quad (2.9)$$

where a is a parameter that determines the fraction of exact exchange included.⁴⁵ This blending leverages the strengths of both Hartree-Fock and DFT, providing a more accurate description of exchange and correlation effects, particularly for systems with significant non-local exchange interactions.

One of the most widely used hybrid functionals is B3LYP, which stands for Becke, three-parameter, Lee-Yang-Parr. The B3LYP functional combines the exact exchange energy with LDA and GGA components, offering a balance between computational efficiency and accuracy. It has been extensively tested and validated for a wide range of molecular systems, often providing superior results for properties such as bond lengths, reaction energies, and activation barriers compared to pure functionals.^{46,47}

The success of hybrid functionals can be attributed to their ability to partially correct the self-interaction error inherent in pure DFT functionals. This error arises because approximate exchange-correlation functionals do not fully cancel the self-interaction of electrons, leading to inaccuracies in the calculated energies. By incorporating a portion of exact exchange, hybrid functionals reduce this error, resulting in more reliable and accurate predictions for both ground-state properties and reaction mechanisms.⁴⁸

In summary, hybrid functionals represent a significant advancement in the field of DFT, offering improved accuracy over pure functionals by incorporating elements of exact exchange from Hartree-Fock theory. Their ability to better account for electron correlation and exchange interactions makes them indispensable tools for the study of complex molecular systems and materials, bridging the gap between computational feasibility and precise modeling of electronic structures.

2.3 Implementing DFT using VASP and Quantum Espresso

2.3.1 Plane-Wave DFT. Plane-wave Density Functional Theory (DFT) is a widely used computational approach in quantum chemistry and materials science for studying the electronic properties of periodic systems, such as crystals and surfaces. The plane-wave basis set, which consists of sinusoidal functions, offers several advantages in terms of computational efficiency and simplicity. In this approach, the electronic wavefunctions are expanded in terms of plane waves, which are particularly well-suited for periodic boundary conditions.⁴⁹

In plane-wave DFT, the wavefunctions $\psi_{n\mathbf{k}}(\mathbf{r})$ of the electrons are expressed as a linear combination of plane waves:

$$\psi_{n\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{G}} c_n^{\mathbf{k}\mathbf{G}} e^{i(\mathbf{k}+\mathbf{G})\cdot\mathbf{r}}, \quad (2.10)$$

where \mathbf{G} are the reciprocal lattice vectors of the form $n_a \mathbf{a}^* + n_b \mathbf{b}^* + n_c \mathbf{c}^*$ (where n_a , n_b , and n_c are integers), \mathbf{k} is the wave vector within the first Brillouin zone, and $c_n^{\mathbf{kG}}$ are the expansion coefficients. This is essentially a Fourier series expansion of the Bloch functions, which take advantage of the periodicity of the crystal lattice, and the size of the basis set is typically determined by a constraint on the maximum total energy E_{\max} of the plane waves (i.e. $\frac{1}{2}|\mathbf{k} + \mathbf{G}|^2 < E_{\max}$).⁵⁰

One of the key benefits of using plane waves in DFT calculations is their ability to systematically converge to the exact solution by increasing the number of basis functions. Plane waves are unbiased and do not favor any particular region of space, making them ideal for describing the delocalized nature of electronic states in solids. Moreover, the use of Fast Fourier Transforms (FFT) allows for efficient computation of the kinetic energy and the application of pseudopotentials, which are used to simplify the treatment of core electrons by replacing them with an effective potential.^{51,52}

Due to its convenient implementation for systems with periodic boundary conditions, plane-wave DFT is a valuable tool for investigating the properties of semiconductors, metals, and insulators, as well as more complex phenomena such as surface states, defects, and adsorption processes.⁵⁰ Despite these advantages, plane-wave DFT also has some limitations. The use of plane waves requires the system to be periodic, which can be challenging for studying isolated molecules or systems with large vacuum regions. However, techniques such as the use of supercells or the inclusion of corrective terms can mitigate these issues. Additionally, the computational cost of plane-wave DFT can be high for systems with very large unit cells or those requiring high precision, but ongoing advancements in computational power and algorithms continue to expand its applicability⁵².

2.3.2 VASP: The State of the Art Periodic DFT Software.

The Vienna Ab initio Simulation Package (VASP) is a highly efficient software package used for performing ab initio quantum chemistry calculations using pseudopotentials and a plane-wave basis set. It is particularly well-suited for the study of solid-state materials and their properties. VASP implements Density Functional Theory (DFT) and allows for the use of various exchange-correlation functionals, including Local Density Approximation (LDA), Generalized Gradient Approximation (GGA), and hybrid functionals.

To start a VASP calculation, one must prepare a few essential input files: the INCAR, POSCAR, POTCAR, and KPOINTS files. While not explicitly stated in the documentation, the naming system for VASP files has a simple structure that is easy to remember. The "CAR" suffix stands for "carrier". So for example the INCAR file stands for "input carrier", as it contains the input parameters that control the calculation (such as the type of exchange-correlation functional to be used, the energy cutoff for the plane-wave basis set, and the convergence criteria). For example, to use the GGA functional by Perdew-Burke-Ernzerhof (PBE), one would set 'GGA = PE' in the INCAR file.⁴³

The POSCAR file defines the initial positions of the atoms in the simulation cell. This file specifies the lattice vectors and atomic coordinates, which can be given in either Cartesian or fractional coordinates. The POTCAR file contains the pseudopotentials for the elements present in the simulation. Pseudopotentials are crucial in VASP as they replace the core electrons with an effective potential, reducing the computational cost significantly while accurately representing the interaction between valence electrons and the atomic cores.⁵¹ You can obtain these

pseudopotentials from the VASP pseudopotential library, ensuring that they are compatible with the chosen exchange-correlation functional.

Finally, the KPOINTS file specifies the k-point sampling of the Brillouin zone. Proper k-point sampling is essential for accurately calculating the electronic properties of periodic systems. For example, a denser k-point grid is needed for metallic systems compared to insulators or semiconductors to capture the electronic structure accurately.⁵⁰

Once the input files are prepared, VASP can be run to perform the calculation. VASP uses an iterative approach to solve the Kohn-Sham equations, starting with an initial guess for the wavefunctions and electron density and refining these in each iteration until self-consistency is achieved. The plane-wave basis set is particularly advantageous in this context because it allows for systematic convergence by increasing the energy cutoff.⁴⁹ During the calculation, VASP utilizes Fast Fourier Transforms (FFT) to efficiently compute the kinetic energy and apply the pseudopotentials.

After the calculation is complete, VASP generates several output files containing the results, such as the total energy, electronic density of states, and optimized atomic positions. These results can be analyzed to gain insights into the material's properties, such as its electronic structure, stability, and response to external perturbations.

2.3.3 Quantum Espresso: An Open Source DFT Package.

Quantum ESPRESSO is a widely used open-source suite for electronic-structure calculations and materials modeling at the nanoscale. It is based on Density Functional Theory (DFT), plane waves, and pseudopotentials, providing a comprehensive set of tools for the simulation of solid-state, liquid, and molecular

systems. Quantum ESPRESSO stands out due to its modular architecture, allowing for flexibility and extensibility in various types of calculations, from ground-state energy computations to more advanced properties such as phonon dispersion and electron-phonon interactions.⁵³

To perform calculations with Quantum ESPRESSO, several input files are required, including the ‘pw.x’ input file for the Plane-Wave Self-Consistent Field (PWscf) calculations. This file specifies the parameters of the calculation, such as the choice of exchange-correlation functional, energy cutoff for the plane-wave basis set, and convergence criteria. For example, the PBE GGA functional can be selected by setting `input.dft='PBE'` in the input file. The atomic structure is described in a separate file, typically in the form of a .xyz or .pdb file, while the pseudopotentials are provided in UPF format. Quantum ESPRESSO supports a variety of pseudopotentials, including norm-conserving, ultrasoft, and projector-augmented wave (PAW) pseudopotentials, which can be obtained from online repositories or generated using specific tools within the suite. The k-point sampling is defined in the ‘K_POINTS’ section of the input file, ensuring accurate integration over the Brillouin zone.⁵⁴

One of the key strengths of Quantum ESPRESSO is its extensive support for parallel computing, which allows users to perform large-scale simulations efficiently. The package employs Message Passing Interface (MPI) for distributed-memory parallelism and OpenMP for shared-memory parallelism, enabling high performance on various computing architectures. Additionally, Quantum ESPRESSO integrates seamlessly with other software packages, such as Wannier90 for Wannier function calculations and Yambo for many-body perturbation theory⁵⁴.

Comparing Quantum ESPRESSO with VASP, both packages are highly efficient and widely used for plane-wave DFT calculations, but they cater to slightly different user needs and preferences. VASP is a commercial software known for its robustness, user-friendly interface, and extensive pseudopotential library, making it a popular choice for industrial applications and academic research. Quantum ESPRESSO, being open-source, offers more flexibility and customization options, which is particularly appealing to users who need to adapt the code for specific purposes or integrate it with other tools.

In terms of functionality, both VASP and Quantum ESPRESSO support a wide range of exchange-correlation functionals, including LDA, GGA, and hybrid functionals. However, VASP often provides better documentation and support, which can be beneficial for new users. On the other hand, Quantum ESPRESSO's open-source nature and active user community provide a collaborative environment where users can contribute to the development and improvement of the software. Additionally, Quantum ESPRESSO's emphasis on parallel computing and its modular structure make it highly scalable for large-scale simulations on high-performance computing clusters.^{49,53}

In summary, both VASP and Quantum ESPRESSO are powerful tools for plane-wave DFT calculations, each with its own strengths. VASP's user-friendly interface and comprehensive pseudopotential library make it a strong choice for many users, while Quantum ESPRESSO's open-source flexibility and scalability are ideal for those who require customization and large-scale parallel simulations.

2.4 Excited State Electronic Structure Methods

2.4.1 Many-Body Perturbation Theory. Many-body perturbation theory (MBPT) is a powerful framework for studying excited state properties in

quantum chemistry, particularly for calculating fundamental and optical band gaps of molecular materials. MBPT extends the Hartree-Fock (HF) approximation by systematically including electron-electron interaction effects beyond the mean-field level. The theory starts from a non-interacting reference system and introduces perturbative corrections to account for electron correlation effects.⁵⁵

In the context of MBPT, the electronic structure problem is often formulated using Green's functions, which describe the propagation of electrons and holes in the system. The central quantity in this approach is the one-electron Green's function, $G(\mathbf{r}, \mathbf{r}'; \omega)$, which can be expressed in terms of the self-energy $\Sigma(\mathbf{r}, \mathbf{r}'; \omega)$ as follows:

$$G(\mathbf{r}, \mathbf{r}'; \omega) = G_0(\mathbf{r}, \mathbf{r}'; \omega) + G_0(\mathbf{r}, \mathbf{r}'; \omega)\Sigma(\mathbf{r}, \mathbf{r}'; \omega)G(\mathbf{r}, \mathbf{r}'; \omega), \quad (2.11)$$

where G_0 is the Green's function of the non-interacting system. The self-energy Σ encapsulates all the many-body effects and can be approximated using different levels of perturbation theory.⁵⁶

MBPT can be connected to Time-Independent Perturbation Theory, where the self-energy Σ acts as a perturbation to the zeroth-order Hamiltonian H_0 . In this framework, the total Hamiltonian H is expressed as:

$$H = H_0 + \Sigma, \quad (2.12)$$

where H_0 represents the Hamiltonian of the non-interacting system and Σ accounts for the electron-electron interactions. This approach provides a systematic way to include correlation effects by treating Σ as a perturbation and calculating its impact on the electronic structure. The perturbative expansion in MBPT is typically carried out in powers of the interaction strength. At each order of perturbation, more complex diagrams representing electron-electron interactions are

included. This systematic inclusion of higher-order interactions allows for a more accurate description of the electronic structure compared to mean-field methods.

MBPT also has a strong connection to Density Functional Theory (DFT), another widely used method for electronic structure calculations. DFT provides a mean-field description of the electronic ground state by solving the Kohn-Sham equations, which map the interacting electron problem onto a system of non-interacting electrons moving in an effective potential. While DFT is highly successful in predicting ground state properties, it often falls short in accurately describing excited states due to its approximate treatment of exchange-correlation effects.⁵⁷

MBPT can be used to improve upon DFT results by providing a more accurate treatment of electron correlation. In practice, DFT can be employed to generate the initial electronic structure, which is then refined using MBPT to include many-body effects. This combined approach leverages the strengths of both methods: the computational efficiency of DFT and the accuracy of MBPT in capturing excited state properties.

By providing a framework to systematically include many-body effects, MBPT forms the foundation for advanced electronic structure methods that go beyond simple approximations. Its applications extend to various fields, including condensed matter physics, quantum chemistry, and materials science, where understanding the detailed electronic structure is crucial for interpreting and predicting material properties.

2.4.2 The *GW* Approximation. The *GW* approximation is a significant advancement in MBPT for studying excited state properties in quantum chemistry. To understand the *GW* approximation, it is essential to start with the

Schrödinger equation and how it relates to the Green's function, which provides an alternative formulation using complex analysis.

The time-independent Schrödinger equation for an electron in a potential $V(\mathbf{r})$ is given by:

$$H\psi_n(\mathbf{r}) = E_n\psi_n(\mathbf{r}), \quad (2.13)$$

where H is the Hamiltonian, $\psi_n(\mathbf{r})$ are the eigenfunctions, and E_n are the corresponding eigenvalues. The Hamiltonian H is typically expressed as:

$$H = -\frac{\hbar^2}{2m}\nabla^2 + V(\mathbf{r}). \quad (2.14)$$

To connect this to the Green's function, we consider the operator $(\omega I - H)$, where I is the identity operator. The Green's function $G(\omega)$ is defined as the inverse of this operator:

$$G(\omega) = (\omega I - H + i\eta)^{-1} = \sum_n \frac{|\psi_n\rangle\langle\psi_n|}{\omega - E_n + i\eta}, \quad (2.15)$$

where ω is the complex frequency and η is an infinitesimally small positive number that ensures causality (i.e. its sign indicates the direction of integration in the complex plane). This operator inversion is crucial because it transforms the problem into one that can be analyzed using the tools of complex analysis. Typically it is written in realspace in the form $G(\mathbf{r}, \mathbf{r}'; \omega) = \langle \mathbf{r} | G(\omega) | \mathbf{r}' \rangle$. The eigenvalues of the Hamiltonian correspond to the poles of the Green's function⁵⁵.

The Green's function is also related to the projection operator P , which projects the state onto the eigenstates of the Hamiltonian, by the following identity:

$$P = \frac{1}{2\pi i} \oint G(\omega) d\omega. \quad (2.16)$$

This form is convenient, as it allows one to represent operators in terms of the eigenstates of the Hamiltonian via a convolution with the Green's function.

The well-known spectral representation of the Green's function is given by:

$$G(\mathbf{r}, \mathbf{r}'; \omega) = \sum_n \frac{\psi_n(\mathbf{r})\psi_n^*(\mathbf{r}')}{\omega - E_n + i\eta}, \quad (2.17)$$

where $\psi_n(\mathbf{r})$ and E_n are the eigenfunctions and eigenvalues of the Hamiltonian, respectively. This relationship highlights how the Green's function acts as a bridge between Schrödinger's eigenvalue problem and complex analysis.

In many-body systems, the Green's function must account for electron-electron interactions. This is accounted for by the self-energy $\Sigma(\mathbf{r}, \mathbf{r}'; \omega)$ in Equation 2.11. This equation expresses the Green's function in terms of the zero-order basis functions (e.g. the Kohn-Sham wavefunctions from a DFT calculation), and can be iterated to the desired level of perturbation theory.⁵⁶

The GW approximation specifically approximates the self-energy Σ as the product of the Green's function G and the screened Coulomb interaction W . This approximation is named after the terms G and W in the self-energy expression:

$$\Sigma(\mathbf{r}, \mathbf{r}'; \omega) \approx i \int \frac{d\omega'}{2\pi} G(\mathbf{r}, \mathbf{r}'; \omega + \omega') W(\mathbf{r}, \mathbf{r}'; \omega'). \quad (2.18)$$

The screened Coulomb interaction W is defined as:

$$W(\mathbf{r}, \mathbf{r}'; \omega) = \int d\mathbf{r}'' \epsilon^{-1}(\mathbf{r}, \mathbf{r}''; \omega) v(\mathbf{r}''), \quad (2.19)$$

where $v(\mathbf{r})$ is the bare Coulomb potential and ϵ^{-1} is the inverse dielectric function that accounts for screening effects⁵⁷.

To calculate the fundamental band gap using the GW approximation, one typically starts with a Density Functional Theory (DFT) calculation to obtain the ground state electronic structure. DFT provides an initial approximation of the electronic states by solving the Kohn-Sham equations, which include an exchange-correlation functional to approximate many-body effects. However, the band gap obtained from DFT is often underestimated due to the approximate nature of

the exchange-correlation functional. To improve this, the GW approximation is employed. In this method, the contribution to the energy from the DFT exchange-correlation functional is subtracted from the Kohn-Sham eigenvalues. The remaining part is then corrected by adding the self-energy terms calculated within the GW framework. Specifically, the self-energy Σ replaces the exchange-correlation potential V_{xc} from DFT. The corrected quasi-particle energies E_n^{QP} are obtained from the equation:

$$E_n^{QP} = E_n^{KS} + \langle \psi_n^{KS} | \Sigma(\omega = E_n^{QP}) - V_{xc} | \psi_n^{KS} \rangle, \quad (2.20)$$

where E_n^{KS} and ψ_n^{KS} are the Kohn-Sham eigenvalues and eigenfunctions, respectively. The self-energy Σ is calculated within the GW approximation, accounting for dynamic screening effects. The difference between the quasi-particle energies of the conduction band minimum and the valence band maximum gives the fundamental band gap. This procedure ensures a more accurate prediction of the band gap, incorporating many-body interactions that are inadequately treated by standard DFT methods.

The GW approximation is particularly effective for calculating quasi-particle energies, which provide a more accurate description of electron addition and removal processes compared to Hartree-Fock or Density Functional Theory (DFT) methods. By incorporating dynamic screening effects, the GW approximation captures essential many-body interactions that are critical for accurately predicting fundamental band gaps and excitation energies in molecular materials.

2.4.3 The Bethe-Salpeter Equation. After obtaining the quasi-particle energies and wavefunctions from a GW calculation, the next step to accurately predict the optical properties of a material, including the optical band gap, involves solving the Bethe-Salpeter equation (BSE). The BSE is a powerful

tool that explicitly includes electron-hole interactions, which are crucial for accurately describing optical excitations.⁵⁵

The BSE starts from the two-particle Green's function, which describes the propagation of an electron-hole pair. This two-particle Green's function L can be expressed in terms of the non-interacting two-particle Green's function L_0 and the electron-hole interaction kernel K . The BSE is given by:

$$L = L_0 + L_0 K L, \quad (2.21)$$

where L_0 is the product of the independent quasi-particle electron and hole states, and the Bethe-Salpeter kernel K includes both the direct and exchange interactions between the electron and the hole.

To solve the BSE, we first need the quasi-particle energies and wavefunctions obtained from the GW calculations. These provide the necessary input for the non-interacting two-particle Green's function L_0 , which is constructed from the product of the quasi-particle electron and hole states. The interaction kernel K is then constructed, typically consisting of the direct term, which accounts for the Coulomb attraction between the electron and the hole, and the exchange term, which includes the exchange interaction.⁵⁶

The BSE can be recast as an eigenvalue problem, where solving the BSE amounts to diagonalizing the two-particle Hamiltonian H_{eh} :

$$H_{eh} \Psi_\lambda = E_\lambda \Psi_\lambda, \quad (2.22)$$

where H_{eh} is the two-particle Hamiltonian, E_λ are the excitation energies, and Ψ_λ are the corresponding exciton wavefunctions. The two-particle Hamiltonian H_{eh} is given by:

$$H_{eh} = H_e + H_h + K, \quad (2.23)$$

where H_e and H_h are the single-particle Hamiltonians for the electron and hole, respectively, and K is the electron-hole interaction kernel.

By solving this eigenvalue problem, we obtain the excitation energies and the corresponding exciton wavefunctions. The lowest excitation energy from the BSE solution corresponds to the optical band gap of the material. This optical band gap is typically smaller than the fundamental band gap calculated from GW due to the binding energy of the exciton, which is the electron-hole pair bound by Coulomb attraction.

The BSE thus provides a detailed and accurate description of the optical properties, capturing excitonic effects that are essential for understanding the absorption spectrum and other optical phenomena in materials. By combining GW and BSE, we achieve a comprehensive and reliable method for predicting both the fundamental and optical band gaps, offering valuable insights into the electronic and optical behaviors of molecular materials.

CHAPTER III

CHEMICAL POTENTIALS

3.1 Overview: Chemical Potentials for Molecular Materials

Accounting for defects has become an integral part of solid-state materials design, as imperfections in repeating lattices can imbue functional emergent properties.^{1,58-61} First-principles calculations have been proven to be a reliable way to study and predict the behavior of these defects,^{62,63} alleviating the challenges of synthesis and characterization of real-world materials, and enabling a targeted effort to realize next-generation atomic assemblies.^{12,64-67} However, the concurrent interest in both computer-aided materials discovery and the emergence of molecular materials has highlighted a new and challenging problem in structure-property prediction for materials composed of molecules. Metal-organic frameworks (MOFs), covalent organic frameworks (COFs) and lower dimensional molecular crystals may have physical properties arising from native defects, but their possible defects are “exotic” because the atoms may be partaking in covalent interatomic bonding.

Generally, first-principles calculations can be used to determine the energy of a balanced chemical reaction where the bulk, non-defective material, $E_{\text{Total}}^{\text{DFT}}(\text{Bulk})$, may host an interstitial or vacancy, taken from or given to some competing phase, $\sum_i n_i \mu_i$, yielding the defective material, $E_{\text{Total}}^{\text{DFT}}(\text{D}^q)$. The defect formation energy, $E_f(\text{D}^q)$, also depends on the charge of the defect and the position of the Fermi level, q and E_F , and requires some computational correctional terms to account for fictitious interactions imposed by limited model sizes, E_{corr} . Together, this is expressed by Eq. 1.1.

Covalent bonds are partially accounted for in the traditional defect formalism given by Eq. 1.1, where the defect formation energy is computed as an

equilibrium reaction that implicitly accounts for all phases of materials available from the nominal composition in the reaction. However, the most challenging aspect of defect calculations is the proper treatment of $\sum_i n_i \mu_i$, the competing phases the defect may form instead of being incorporated into the material. In conventional solid-state ionic applications, $\sum_i n_i \mu_i$ is treated by one of two strategies; i) declare the use of a “synthetically-relevant” environment in which the defect may form (*e.g.*, iodine-poor conditions to predict I-vacancy formation energies in halide perovskites⁶⁸) or, ii) exhaustively sample the energies of all competing phases with the nominal compositions of the atoms within the system⁶⁹ (*e.g.*, in adatoms in BaSnO₃, compare the formation enthalpies of all known materials of nominal composition Ba_xSn_yO_z and combinations thereof with the interstitial adatom.⁷⁰). In the former, only one I-containing phase is computed (*e.g.* I₂), whereas the latter required 45 unique materials calculations.

For molecular materials constructed from organic building blocks, the treatment of $\sum_i n_i \mu_i$ poses new challenges because there are seemingly endless combinations of C, H, O, and N, making exhaustive sampling of all possible competing phases an impossible task. This paper discusses how we should treat defects in molecular materials when building blocks are collections of covalently bound atoms, rather than single point charges in ionic lattices, and draws specific attention to materials with organic motifs. We will demonstrate that the strategy of selecting synthetically-relevant phases is difficult but necessary to enable predictive calculations defects in molecular materials.

3.2 Crystalline tetrahydrofuran with interstitial tetrahydrothiophene: an O_S substitution?

To first illustrate the problem associated with treating molecular materials as a kinetically trapped phase on a potential energy surface comprised of all organic molecules, we can examine the simple case of S-substitution in crystalline tetrahydrofuran (THF). In effect, the substitution is a point defect where the resultant material contains a single tetrahydrothiophene (THT) in a crystalline sea of THF.⁷¹ Like all substitution defects, swapping O for S results in a local reorganization of the nearby atoms, associated with differences in electrostatic potential, orbital overlap, and charge density. The same is true here, but with two additional considerations. First, THT and THF crystallize in different space groups, in part because the S lone pair has a larger radius due to breaking kinosymmetry. In materials with weaker bonding (*e.g.* ionic or soft covalent), the effect is not as strongly observed because of the reduced impact of sulfur’s radial node. Second, the effect of S and O substitutions is most pronounced for p-based σ bonding systems, where large energy splitting is observed for bonding and antibonding molecular orbitals. Far fewer non-carbon materials held together with a p-based σ bonds than there are organic molecules composed therefrom.

The THT-in-THF defect can be computed in two conceptually dissimilar ways. In the former, we can treat the defect as an S atom occupying an O site;

$$E_f(\text{SO}) = E_{\text{Total}}^{\text{DFT}}(\text{SO}) - E_{\text{Total}}^{\text{DFT}}(\text{THF}) + \mu_{\text{O}} - \mu_{\text{S}}, \quad (3.1)$$

where the chemical potentials μ_{O} and μ_{S} are referenced to the energies of a single atom in the O₂ and S₈ molecules, respectively. This approach is an O-rich S-rich potential, and should result in the highest concentration of S-atoms incorporated in the THF crystal. Yet, the formation enthalpy is predicted to be large, 2.19 eV. This

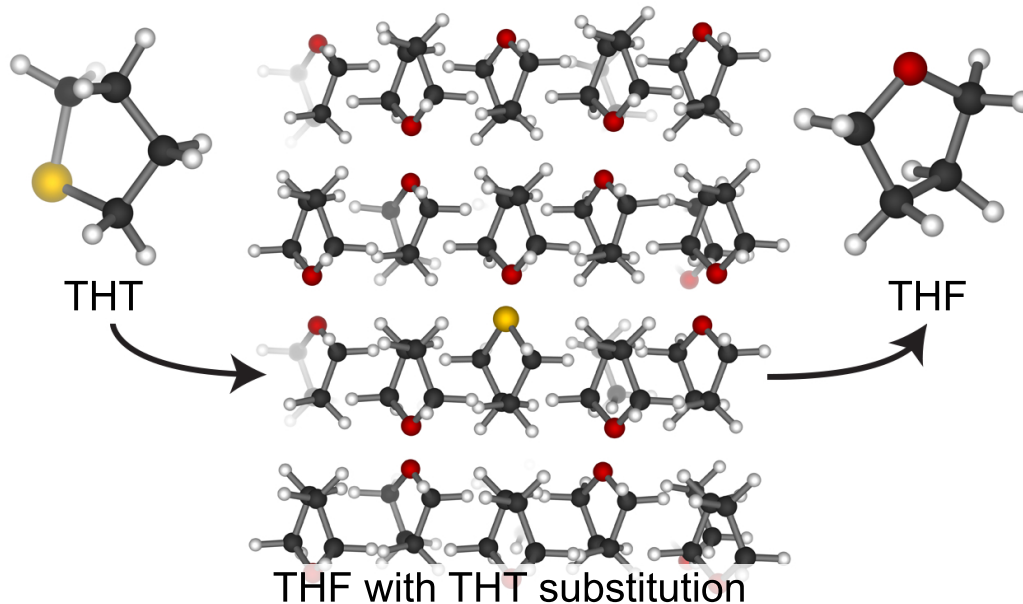


Figure 5. Schematic of a THT molecule replacing a THF molecule in a THF molecular crystal

suggests that either S_8 is stabilizing, O_2 is destabilizing, or that THT in the THF lattice is highly unfavored. The latter would seem unlikely, however, because the molecules have similar dipole moments,⁷¹ and a comparison of their crystal packing factors indicates that the distortion in the crystal is negligible (0.5% difference). Instead, the effect is likely driven by the fact that S-S bonds are weaker and longer (*ca.* 215 kJ/mol, 2.0 pm) than C-S bonds (*ca.* 260 kJ/mol, 1.8 Å), but O=O bonds are stronger and shorter (*ca.* 498 kJ/mol, 1.2 Å) than C-O bonds (*ca.* 350 kJ/mol, 1.4 Å).⁷²

If we instead treat the defect as a molecule rather than an atom by;

$$E_f(\text{THT}_{\text{THF}}) = E_{\text{Total}}^{\text{DFT}}(\text{THT}_{\text{THF}}) - E_{\text{Total}}^{\text{DFT}}(\text{THF}) + \mu_{\text{THF}} - \mu_{\text{THT}} \quad (3.2)$$

where the chemical potentials of THF and THT are each referenced to the energy of the single THF and THT molecules. Doing so yields a formation energy of 18

meV (which corresponds to roughly 2% substitution in the crystal), suggesting that it is not the inclusion of THT in THF that is dictating the 2.19 eV formation energy, but rather due to an improper penalty for cleaving O₂.

The opposite case (i.e. THF substitution in a THT crystal) was also considered. However, in this case using elemental oxygen and sulfur as chemical potential references yielded a defect formation energy of -2.06 eV (which would indicate that the substitution is more stable than the pristine crystal). We explored this ostensibly aphysical result by examining both 50% substituted and fully substituted system (i.e. systems where half or all THT molecules in the crystal were replaced with THF). Phonon calculations on the optimized structure for both cases yielded 24 imaginary frequencies, thus indicating unstable structures and supporting our hypothesis that treating the defects as atomic yields an aphysical result.

On the other hand, when treating the defect as a molecular substitution, the formation energy was calculated to be 60 meV (corresponding to a roughly 10% substitution). This result is much more reasonable - not only is the formation energy positive, but it also indicates that the substitution is more stable than the THT-in-THF substitution. Considering the large difference in size between these two molecules, it makes sense that substituting THT for a smaller THF molecule would be more favored due to decreased steric interactions.

3.3 Defects from Molecular Sources

3.3.1 Hydrogen from a molecular source. We will now consider the case of an atomic defect in a molecular material, where the source of the defect is a molecule. To better understand why this distinction is important, consider the reaction of some bulk material with some hydrogen-containing molecule HX that

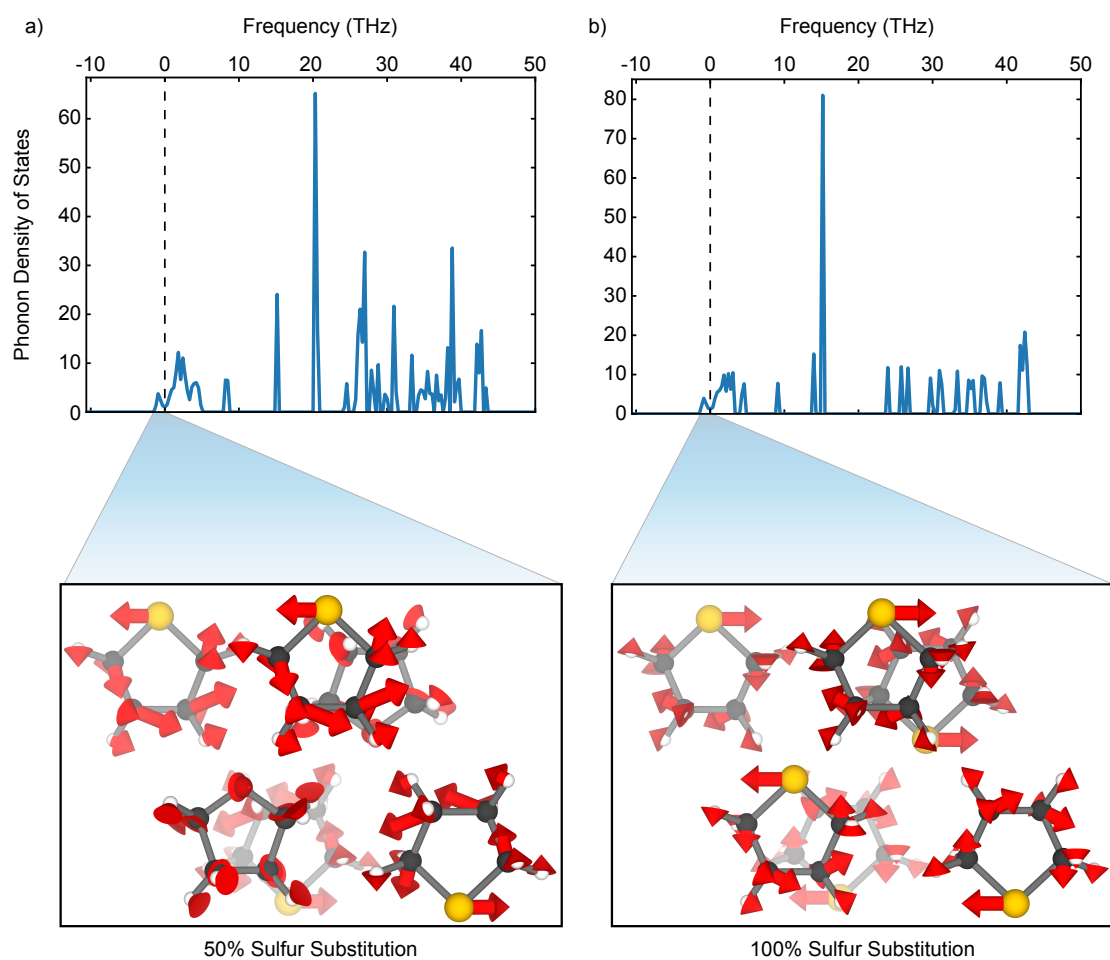
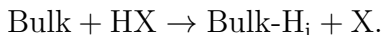


Figure 6. A single imaginary frequency at the Γ -point corresponding to an unoptimized optical phonon mode in crystalized THF with 50% substitution of THT. Such an optical mode did not show up in pristine THF, pristine THT, or in THT with 50% substitution of THF, indicating that 50% substitution of THT in the THF lattice is not dynamically stable.

readily donates a hydrogen atom. If this hydrogen atom goes to form a neutral interstitial defect (H_i) in the material, the chemical reaction is



If we compute the energy of each species via DFT, the formation energy for this reaction is:

$$E_f(\text{Bulk-}H_i) = E_{\text{Total}}^{\text{DFT}}(\text{Bulk-}H_i) + E_{\text{Total}}^{\text{DFT}}(\text{X}) - E_{\text{Total}}^{\text{DFT}}(\text{Bulk}) - E_{\text{Total}}^{\text{DFT}}(\text{HX}). \quad (3.3)$$

However, if we compute the formation energy of the *neutral* H_i defect using Equation 1.1, we have that

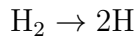
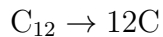
$$E_f(\text{Bulk-}H_i) = E_{\text{Total}}^{\text{DFT}}(\text{Bulk-}H_i) - E_{\text{Total}}^{\text{DFT}}(\text{Bulk}) - \mu_H, \quad (3.4)$$

which seems to contradict the previous equation. This can easily be reconciled if the chemical potential of hydrogen is equal to the change in energy of the hydrogen donor molecule upon losing the hydrogen atom (i.e. $\mu_H = E_{\text{Total}}^{\text{DFT}}(\text{HX}) - E_{\text{Total}}^{\text{DFT}}(\text{X})$), but this choice ostensibly opposes the current defect literature and must therefore be carefully justified.

The standard practice in the literature for defect chemical potentials is to construct them stoichiometrically using Hess’s Law to reference the energies of each atomic species in a defect source molecule⁸. This is often helpful because Hess’s Law allows one to construct a balanced chemical reaction using an external reference energy for the defect source. For computations this is particularly useful when studying systems with unknown absolute energies. This method, in principle, would work for a molecule like benzene whose hydrogens are all in the same chemical environment. However, for more complex molecules it becomes obvious that treating all hydrogens the same is inadequate. Consider for example the molecule 9,10-dihydroanthracene (DHA), which has the chemical formula $C_{14}H_{12}$.

While this molecule has 12 C-H bonds, a trained chemist would see that some of these bonds would be easier to break than others. Therefore, one would anticipate that certain hydrogen atoms on this molecule would be easier to remove than others. We will show that the literature method for this case fails to adequately construct a reasonable chemical potential for hydrogen, and argue that this is because it fails to account for the changes in the overall potential energy landscape by treating the hydrogen atoms as chemically and stoichiometrically equivalent.

Using the equivalent atoms approach as in ref.⁸, we can construct the energy of a single hydrogen atom from DHA by writing



where C_{12} represents a unit cell of graphite (the most stable elemental phase of carbon). These gives the following chemical potential relationships:

$$6\mu_{\text{DHA}} = 7\mu_{\text{C}_{12}} + 36\mu_{\text{H}_2} \quad (3.5)$$

$$\mu_{\text{C}_{12}} = 12\mu_{\text{C}} \quad (3.6)$$

$$\mu_{\text{H}_2} = 2\mu_{\text{H}} \quad (3.7)$$

Solving for the chemical potential of hydrogen gives that $\mu_{\text{H}} = \frac{1}{12} (\mu_{\text{DHA}} - 14\mu_{\text{C}})$.

Since we want to reference μ_{H} to the DHA molecule rather than elemental hydrogen, we compute the energy of a DHA molecule and the energy of a graphite unit cell using DFT. Thus, we arrive at the chemical potential:

$$\mu_{\text{H}} = \frac{1}{12} \left[E_{\text{Total}}^{\text{DFT}}(\text{DHA}) - \frac{7}{6} E_{\text{Total}}^{\text{DFT}}(\text{graphite}) \right] \quad (3.8)$$

The factor of $1/12$ in Equation 3.8 explicitly demonstrates that all twelve hydrogen atoms in the molecule are being treated equally. However, it is known that the hydrogens at the 9 and 10 position on DHA are about 20% weaker than typical C-H bonds⁷³. This is because dehydrogenation at the 9 and 10 positions restores aromaticity to the system, yielding a downhill potential energy gradient. Therefore, if DHA is the source of an H_i defect in a material, one should anticipate the following reaction to occur:



with formation energy

$$E_f(H_i) = E_{\text{Total}}^{\text{DFT}}(\text{Bulk-}H_i) + \frac{1}{2}E_{\text{Total}}^{\text{DFT}}(\text{Anthracene}) - E_{\text{Total}}^{\text{DFT}}(\text{Bulk}) - \frac{1}{2}E_{\text{Total}}^{\text{DFT}}(\text{DHA}). \quad (3.9)$$

Using Equation 3.9, we can write the chemical potential of hydrogen as

$$\mu_H = \frac{1}{2} (E_{\text{Total}}^{\text{DFT}}(\text{DHA}) - E_{\text{Total}}^{\text{DFT}}(\text{Anthracene})). \quad (3.10)$$

This can also be justified by looking at the chemical potential relationship $\mu_{\text{DHA}} = \mu_{\text{Anthracene}} + 2\mu_H$, which gives

$$\mu_H = \frac{1}{2} [\mu_{\text{DHA}} - \mu_{\text{Anthracene}}], \quad (3.11)$$

which lines up perfectly with Equation 3.10. This equation can be understood as a phase relationship between DHA, anthracene, and hydrogen rather than DHA, carbon, and hydrogen. The key difference between this approach and the one shown before it is that here we have explicitly connected the hydrogen atoms being exchanged between two H-containing phases, therefore accounting for the specific chemical behavior of those particular atoms.

Calculating the chemical potential using Equation 3.8 yields $\mu_H = -2.97$ eV. Calculating half the DFT energy of an H_2 molecule gives us the standard reference

chemical potential for hydrogen to be $\mu_{\text{H}}^{\circ} = -3.26$ eV, which is significantly lower in energy. Based on Equation 3.4, this would imply that DHA is a better source of H_i defects than H_2 (therefore exceeding the so-called “H-rich” condition often referenced in the defect literature). If we instead use Equation 3.10, we obtain a value of $\mu_{\text{H}} = -3.71$ eV, which is 450 meV lower in energy than μ_{H}° . Clearly these approaches yield completely different results; to benchmark the utility of both, we will apply this to a defect analysis of the photocatalytic MOF MIL-125 and compare the results.

3.3.2 Hydrogen Defects in a Titanium-Based MOF. MIL-125 is a titanium based MOF that is known to act as a photocatalyst in the presence of primary and secondary alcohols, in a similar priority to oxidation with sodium dichromate as taught in an undergraduate organic chemistry course⁷⁴. As discussed by Mancuso et al.,⁷⁵ upon UV radiation the alcohols are oxidized to form the corresponding aldehyde or ketone, and the two liberated hydrogen atoms redox pair with the Ti (IV) in the secondary-building unit (SBU) to form H_i defects. This results in reduction of the Ti (IV) to Ti (III), causing the solution to turn black.

Our goal here is to look at this process purely from a thermodynamic perspective using Equation 3.4, applying both approaches outlined above to predict which alcohols are better suited for this reaction. In other words, we want to predict which alcohols are better sources of hydrogen for this reaction by calculating the resulting concentration of H_i defects from each one. The purpose of this exercise is to illustrate the stark difference in physicality between the two approaches, and to illustrate that experimental validation is needed to establish a valid reference for atomic chemical potentials. The relative defect concentration is

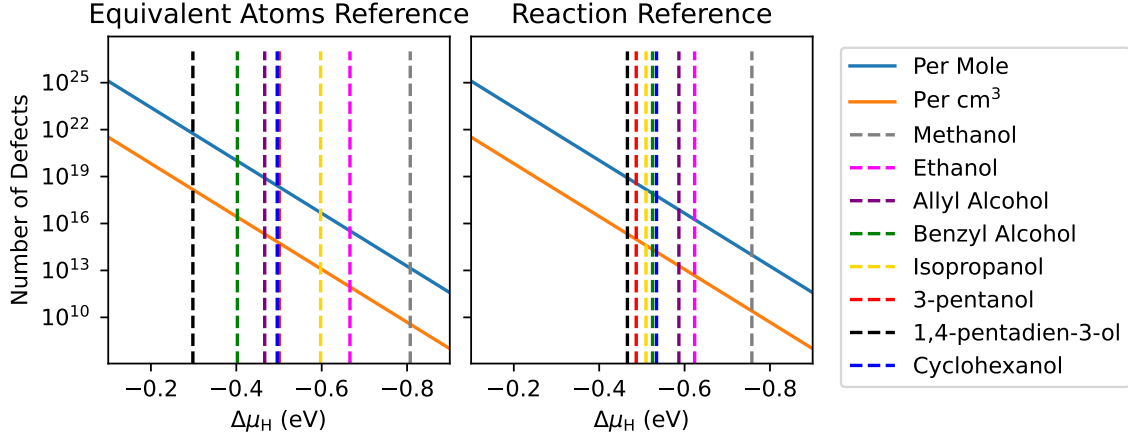


Figure 7. Predicted number of H_i defects in MIL-125 as a function of the (relative) chemical potential of hydrogen. The dashed lines depict chemical potentials calculated from various alcohols using Eqs. 3.13 (left) and 3.14 (right).

calculated as

$$C(H_i) = N \exp \left\{ -\frac{E_f(H_i)}{k_B T} \right\}, \quad (3.12)$$

where N is the number of equivalent defect sites, and $C(H_i)$ is given as a mole fraction with respect to the bulk material^{1,76}. Multiplying Equation 3.12 by Avogadro's number gives the number of defects per mole of the bulk material.

The reason we are modelling this formation process from ground-state MIL-125 (as opposed to modelling the excited state) is that the formation of interstitial hydrogen after excitation is a highly spontaneous process. If we calculate the formation energy of these defects starting from MIL-125 in the triplet state to *simulate* the excited state we see that the reaction is highly exothermic, yielding formations energies as low as -2.96 eV. At this point, we would expect the reaction to depend mainly on kinetics rather than thermodynamics, which is outside the scope of this paper.

Table 1. Comparison of the *ab initio* (absolute) hydrogen chemical potentials referenced to select alcohols using Eqs. 3.13 and 3.14. All values are given in units of eV.

Primary Alcohols	Equal Atoms	Reaction	Difference
Methanol	-4.06	-4.01	0.05
Ethanol	-3.92	-3.88	0.04
Propanol	-3.83	-3.88	-0.05
Isobutanol	-3.79	-3.87	-0.08
Allyl Alcohol	-3.72	-3.84	-0.12
2-phenylalcohol	-3.64	-3.84	-0.20
Benzyl Alcohol	-3.66	-3.78	-0.12
Secondary Alcohols			
Isopropanol	-3.85	-3.77	0.08
2-pentanol	-3.76	-3.76	0.00
3-pentanol	-3.76	-3.74	-0.02
3-buten-2-ol	-3.71	-3.75	-0.04
1,4-pentadien-3-ol	-3.55	-3.72	-0.17
1-phenylethanol	-3.66	-3.72	-0.06
Cyclohexanol	-3.75	-3.79	-0.04

The hydrogen chemical potential for an alcohol with chemical formula C_nH_mOH is computed by the equal-atoms approach as

$$\mu_H = \frac{1}{m+1} \left[E_{\text{Total}}^{\text{DFT}}(C_nH_mOH) - \frac{n}{12} E_{\text{Total}}^{\text{DFT}}(\text{graphite}) - \frac{1}{2} E_{\text{Total}}^{\text{DFT}}(O_2) \right], \quad (3.13)$$

whereas using the reaction discussed above as the reference gives

$$\mu_H = \frac{1}{2} \left[E_{\text{Total}}^{\text{DFT}}(\text{Alcohol}) - E_{\text{Total}}^{\text{DFT}}(\text{Carbonyl}) \right]. \quad (3.14)$$

The results of Equations 3.13 and 3.14 for several selected primary and secondary alcohols are given in Table 1. Figure 7 shows the molar and volumetric defect concentration as a function of μ_H using both methods.

It is worth noting that the approach we employ in Eq. 3.14 does also treat the non-reactive hydrogens in the alcohol (and the resulting carbonyl product) as equivalent. We justify the explicit treatment of the reactive hydrogens differently

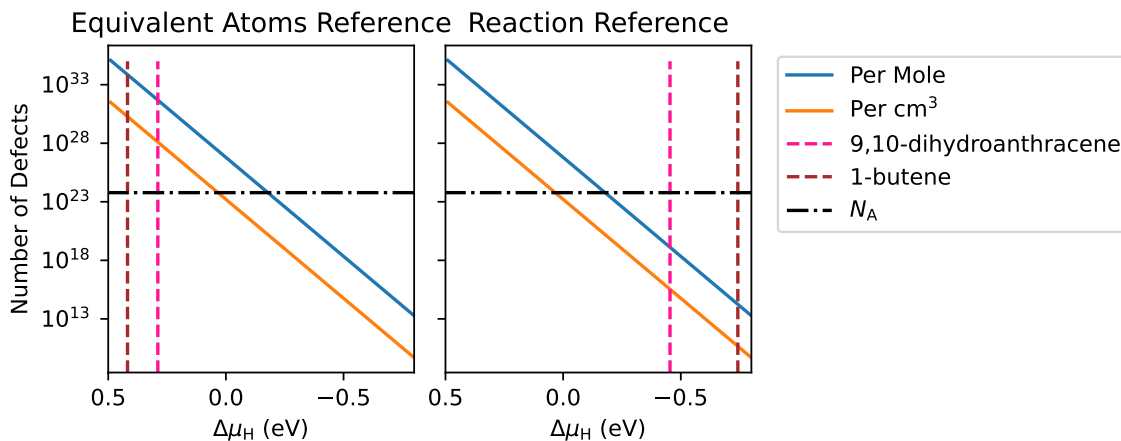


Figure 8. Predicted number of H_i defects in MIL-125 as a function of the (relative) chemical potential of hydrogen. The chemical potentials are calculated from 9,10-DHA and 1-butene using the two schemes discussed in the text.

by calculating their relative pK_a values. Using ethanol as an example, we note that both protons involved in the reaction have higher pK_a s (15.8 for the OH proton and 45.4 for the alkyl proton) than those on the neighboring CH_3 group (which have a pK_a of 93.8). Additionally, the pK_a s for the reacting protons drop significantly if we remove an electron first (to -38.5 and -43.4, respectively), suggesting that this is an electron-first redox event that results in both protons being removed more-or-less simultaneously.

While some alcohols such as methanol showed little variation in μ_H between the two methods, others showed as much as a 200 meV difference. All of the alcohols with significantly different μ_H values had properties that were chemically significant and thermodynamically stabilizing. As a result, the two methods gave completely different defect concentrations (Figure 7). The most notable difference is that Equation 3.14 gave a trend that seems to line up with chemical intuition - for example, molecules that were stabilized via resonance and aromaticity upon oxidation appeared to be better sources of hydrogen.

As a final benchmark, we will pretend that our DHA example from above, and also 1-butene (as the butadiene product would be resonance stabilized), are sources of hydrogen for this photodoping process. The resulting molar and volumetric defect concentrations are plotted in Figure 8. Interestingly, using the equal-atoms approach to compute μ_{H} gives concentrations on the order of 10^{37} - several orders of magnitude larger than Avogadro's number. This is significantly more than a mole of defects per mole of the host material, which would imply that MIL-125 is unstable in the presence of both of these molecules. Using chemical reactions to reference μ_{H} yields more reasonable results, giving concentrations on the order of 10^{15} - 10^{20} (and also reversing the order of μ_{H} referenced to each molecule).

3.4 Discussion

We have presented here an intuitive method for calculating chemical potentials for defective species based on the chemical and physical processes involved in the process for forming the defect. For molecular defects we reference the chemical potential to the total DFT energy of that molecule due to the inseparable nature of the covalent interactions within that species. For atomic defects obtained from a reaction, we account for the thermodynamics of breaking and forming covalent bonds by calculating the energetic difference between the the defect-donating species before and after it has given up the defective atom. Like previous methods, this method requires one to know all the relevant chemical species participating in the reaction, especially when there are multiple potential sources for a defect. In addition, one must know the relevant balanced reactions and *stable* products from the donating species. It is our hope is that experimentalists will test this method for predicting defect concentrations.

CHAPTER IV

BAND GAPS

4.1 Current Methods Underestimate Fundamental Band Gaps of MOFs

Calculating band gaps for molecular materials, particularly metal-organic frameworks (MOFs), presents unique challenges due to the intricate nature of their structures and the interactions within. MOFs consist of metal ions or clusters coordinated to organic ligands, forming a wide range of 2- and 3-dimensional crystalline structures. Many MOFs are known to behave as flat-band insulators due to highly localized electronic states, strong electron correlation, and quantum confinement effects.⁷⁷ A significant obstacle in calculating the band gaps of MOFs stems from the intricate relationship between metal ions and organic ligands. The ionic interactions between these components often lead to decreased conductivity and the generation of localized excitons. In contrast to conventional semiconductors with dispersed electronic states that aid in charge transfer, electronic states in MOFs tend to be more confined due to the strong ionic nature of metal ligand bonds. This confinement results in limited mobility for charge carriers, influencing the material's overall electronic characteristics. Moreover, the hybrid nature of MOFs, combining inorganic and organic components, complicates the use of conventional band structure calculation methods. Density Functional Theory (DFT), often employed for band gap calculations, may require advanced functionals to accurately describe the exchange-correlation effects in MOFs. The presence of localized d or f orbitals in the metal centers further adds to the complexity, necessitating the inclusion of Hubbard corrections (DFT+U) or hybrid functionals to obtain reliable band gap values.⁷⁷

The highly localized excitons, resulting from strong Coulomb interactions within MOFs, pose another significant challenge. These excitons, which are electron-hole pairs bound together by electrostatic forces, tend to be more confined in MOFs compared to traditional semiconductors. This confinement can lead to larger exciton binding energies, resulting in a drastic difference between the optical band gap E_{opt} (defined as the energy needed to create an exciton) and the fundamental band gap E_g (defined as the difference in energies between the highest occupied molecular orbital (HOMO) and the lowest unoccupied molecular orbital (LUMO)). This drastically affects the optical properties of the material and complicates the interpretation of experimental spectroscopic data.^{78,79} In conventional semiconductors E_{opt} and E_g are often close enough to be considered the same. However, if a material has a large exciton binding energy then optical measurements of the band gap are no longer a reliable way to measure the fundamental gap.

Additionally, standard DFT methods such as LDA and GGA are known to underestimate the band gap of materials due to self-interaction errors. This underestimation is particularly problematic in MOFs, where the band gaps are already influenced by the highly localized nature of excitons. The inherent limitations of LDA and GGA in accurately capturing the electronic structure and the strong correlation effects in MOFs compound the issue of localized excitons and strong exciton binding energies. Hybrid functionals have been shown to somewhat correct for this, giving reliable values for E_{opt} , but accurate calculation and experimental measurement of the fundamental band gaps for most MOFs remains to be seen.¹³

The unique structural and electronic characteristics of MOFs, driven by the ionic interactions between metals and organic ligands, present significant challenges in accurately calculating their band gaps. To address these challenges, techniques such as the GW approximation and Bethe-Salpeter equation (BSE) can be employed.⁷⁸ MBPT methods account for electron-electron and electron-hole interactions more accurately than standard DFT, providing improved predictions of electronic excitations and band gaps in complex molecular materials like MOFs.

4.2 Motivation: Previous Studies on Organic-Inorganic Halide Perovskites

Organic-inorganic halide perovskites have garnered significant attention due to their promising applications in optoelectronic devices, particularly solar cells. Accurate band gap predictions are crucial for optimizing these materials and tuning their properties. Many-body perturbation theory (MBPT) has emerged as a powerful tool for this purpose. MBPT has been extensively used to calculate the band gaps of organic-inorganic halide perovskites, providing critical insights into their electronic properties and guiding the design of advanced optoelectronic devices. The success of MBPT in accurately predicting the electronic properties of these complex systems underscores its potential applicability to other similarly intricate materials, such as metal-organic frameworks (MOFs). By leveraging the lessons learned from perovskites, we can extend MBPT methodologies to a broader range of molecular materials, thereby advancing the field of computational materials science.

One of the key successes of MBPT in perovskite research is its ability to provide highly accurate band gap predictions. For instance, Filip et al. (2014) demonstrated the effectiveness of the GW approximation in predicting the

quasiparticle band gaps of hybrid organic-inorganic perovskites.⁸⁰ This level of accuracy is essential for designing materials with optimal optoelectronic properties. Additionally, MBPT has proven invaluable in modeling excitonic effects and electron-hole interactions, as highlighted by Leppert (2024).⁸¹ Her work on exciton binding energies in perovskites provided a deeper understanding of the material's electronic structure, which is critical for enhancing the performance of solar cells. Another recent and significant achievement of MBPT is its ability to elucidate radiative recombination rates. Davies et al. (2016) explored the radiative bimolecular recombination in methylammonium lead triiodide, revealing it to be an inverse absorption process.⁸² This type of understanding is crucial for improving charge-carrier lifetimes and consequently the efficiency of photovoltaic devices. These successes illustrate the utility and capability of MBPT in tackling complex electronic interactions and providing detailed insights into the material properties that are pivotal for device performance.

Despite these successes, applying MBPT to perovskites is not without challenges. One of the primary obstacles is the high computational cost associated with these calculations. Leppert et al. (2019) discussed the significant computational resources required for MBPT, which can be a limiting factor in studying large or complex systems. Additionally, handling structural disorder and phonon interactions in perovskites presents another layer of complexity.^{83,84} Gao et al. (2016) highlighted the difficulties in accurately modeling disordered systems and the impact of structural distortions on electronic properties.⁸³ Umadevi et al. (2019) further investigated this specifically in the context of a lead-free hybrid germanium iodide perovskite.⁸⁴ These challenges must be addressed to fully leverage MBPT's potential in studying perovskites and other molecular materials.

Furthermore, accurately accounting for screening effects and Coulomb interactions remains a challenge in MBPT applications. Vona et al. (2021) provided insights into the complexities of modeling Coulomb screening and choice of functional setting up MBPT calculations for perovskites.⁸⁵ Overcoming these challenges requires advancements in computational methods and a deeper understanding of the underlying physics. Another notable challenge is the accurate inclusion of spin-orbit coupling (SOC) effects, which are significant in heavy elements like lead. Both Filip et al. (2014) and Gao et al. (2016) discussed the impact of SOC on the electronic structure of perovskites, highlighting the necessity to incorporate these effects to achieve accurate predictions.^{80,83} The inclusion of SOC in MBPT calculations increases the computational burden and complexity but is essential for precise modeling of materials with heavy elements.

The lessons learned from applying MBPT to perovskites can be extended to other molecular materials, such as metal-organic frameworks (MOFs). The accurate band gap predictions and detailed insights into electronic interactions achieved in perovskites suggest that MBPT can be equally effective for MOFs and other complex materials. Addressing the computational and modeling challenges in perovskites will pave the way for broader applications of MBPT, enhancing our ability to design and optimize a wide range of molecular materials. The successful application of MBPT to organic-inorganic perovskites has provided valuable insights into their electronic properties and demonstrated the theory's potential for broader applications. While challenges remain, particularly in terms of computational cost and modeling complexity, the advancements made in perovskite research offer a promising foundation for extending MBPT to other molecular materials. By continuing to refine MBPT methodologies and addressing existing

challenges, we can unlock new possibilities in the design and optimization of advanced materials for various optoelectronic applications.

4.3 Many-Body Perturbation Theory as a way to accurately obtain E_g and E_{opt}

4.3.1 Challenges of applying MBPT to MOFs. MBPT is a complex computational method used to account for electron interactions that aren't fully captured by simpler approaches like DFT. Among various MBPT methods, the GW approximation is known for its ability to adjust electronic band structures, making it useful for materials with strong electron correlations such as MOFs. However, it poses challenges due to its high computational requirements. One major hurdle is the significant computational cost associated with calculating the Green's function and self energy, which involves summing over all the system's states, including numerous unoccupied states, for accuracy. This leads to the creation of large matrices that demand substantial memory and computing power, scaling as N^4 with the number of atoms N . As a result, applying GW to very large systems becomes impractical without advanced computing resources or simplifications.^{86,87}

Accurate numerical integration in both reciprocal (k-space) and frequency space plays a crucial role in MBPT calculations.⁸⁸ Achieving precise integration in reciprocal space often necessitates dense k-point sampling to adequately cover the Brillouin zone. In frequency space, the integration involves complex contour integrals, and achieving convergence in these integrals can be particularly challenging due to the need to accurately capture the energy dependencies of the Green's function and the screened Coulomb interaction. Conducting convergence tests for every system is a crucial aspect of MBPT computations. Different systems

often require specific parameters to achieve convergence, such as varying numbers of k points, frequency points and unoccupied states considered in the computations. This need for tailored convergence testing for each system adds complexity and computational requirements since it often entails running multiple calculations with different settings to pinpoint the optimal settings.

MOFs present additional complexity when applying MBPT due to their unique structural and electronic properties.^{78,79} One of the biggest challenges is the size of the unit cells for most MOFs. These unit cells can contain hundreds of atoms, significantly increasing the computational burden compared to simpler materials. The size of the unit cell not only amplifies the N^4 scaling issue but can also necessitates careful consideration of k-point sampling to achieve accurate results to reduce the computational demand.

The dielectric properties of MOFs can also pose a challenge to MBPT calculations. The porosity of MOFs leads to a low and highly inhomogeneous dielectric constant. Accurately converging the dielectric function for such materials requires careful consideration and extensive computational effort. This inhomogeneity arises from the complex structure and variability in the local environment within the framework, necessitating meticulous convergence checks. The low dielectric constant can affect the screening of electron-electron interactions, leading to stronger Coulomb interactions that must be accurately captured by MBPT calculations. The weak screening from the low dielectric constant and the ionic nature of metal-ligand interactions within the framework also lead to strong excitonic effects. The weak screening enhances electron-hole interactions, which in turn necessitates the inclusion of more virtual states to achieve convergence for both the dielectric function and the self-energy. The increased number of

required virtual states further amplifies the computational demands of applying MBPT to MOFs. For example, excitons in MOFs can have binding energies that are significantly larger than those in traditional semiconductors, requiring more detailed and extensive calculations to accurately describe these states.

While MBPT provides a robust framework for capturing electron correlation effects, its application to MOFs is fraught with challenges. The large unit cells, complex dielectric properties, and strong excitonic effects unique to MOFs necessitate careful and computationally intensive approaches to achieve accurate results. Moreover, the chemical diversity and flexibility of MOFs introduce additional layers of complexity. Different metal centers and organic linkers can drastically alter the electronic properties, necessitating system-specific adjustments such as the inclusion of spin-orbit coupling.⁸⁰ The presence of various functional groups within the MOF structure can also lead to localized states that require fine-tuned computational techniques to accurately model their effects on the overall electronic structure. These challenges underline the need for simplifying approximations to expedite calculations, as well as continued advancements in computational methods and resources to fully harness the potential of MBPT in the study of complex materials like MOFs.

4.3.2 Methods to improve computational efficiency. In order to successfully apply MBPT to MOFs, simplifying approximations and techniques must be employed to improve the computational efficiency of these calculations, ideally without sacrificing accuracy. One effective approach to reducing computational costs is to minimize the number of atoms used in these calculations, namely through the use of the primitive unit cell of the MOF whenever possible. The primitive unit cell contains the smallest number of

atoms that retain the full symmetry and periodicity of the crystal structure. By minimizing the number of atoms in the unit cell, we can significantly reduce the computational resources required for MBPT calculations. This reduction not only decreases the size of the matrices involved but also limits the number of k-points needed for accurate Brillouin zone (BZ) sampling. For example, in large MOFs where supercells are often used to model defects or adsorbates, reverting to the primitive unit cell can save substantial computational time while preserving the essential physical properties of the system. This situation is ideal for MBPT calculations on pristine MOF structures, but unfortunately any modification to the MOF that changes its symmetry requires use of the full conventional unit cell.

For flat-band insulating MOFs, such as MOF-5 and MIL-125, using only the Γ point in BZ sampling can also be an effective strategy to reduce computational cost.⁷⁸ Since these materials exhibit minimal band dispersion, their electronic properties (particularly their band gaps) are relatively insensitive to k-point sampling. However, since the screened Coulomb potential diverges near the Γ -point, this approach requires careful numerical integration of the BZ to avoid inaccuracies. One method to achieve this is the Random Integration Method (RIM), which improves the precision of the BZ integration through a random sampling of k-points within the BZ using a three-dimensional Monte Carlo technique.⁸⁹ This method ensures that the Γ point approximation captures the essential electronic properties without the need for a dense k-point grid, significantly reducing computational efforts while maintaining accuracy.

Since the primary cost of MBPT calculations is that they require a large number of unoccupied bands, many attempts at reducing computational cost focus on reducing the number of bands required for convergence. These range

from approximations to minimize sum truncation error to more exact contour deformation techniques.^{90,91} The Bruneval-Gonze (BG) terminator is a technique designed to reduce the number of empty states required in GW calculations by truncating the sums required for calculating the polarizability and self-energy and replacing the remaining terms with a correction, thus lowering computational demands. This method uses an extrapolar approximation to replace terms for high-energy states with terms involving a average "common energy", allowing for an accurate representation of the self-energy with fewer virtual states. By employing this technique, we can maintain the accuracy of the GW calculations while significantly reducing the computational cost. This approach is particularly useful for MOFs, where the inclusion of a large number of virtual states can become prohibitive. By truncating the number of required bands, the BG terminator effectively streamlines the computational process, making MBPT calculations more feasible for large and complex MOF systems.

Another approach to improving computational efficiency is to simplify the GW approximation itself, such as through single-shot GW (G_0W_0). For this method, the Green's function and the screened Coulomb interaction are calculated using a zero-order basis from a non-interacting reference system, such as the Kohn-Sham orbitals obtained from DFT. This single-shot approach provides a good first approximation of the quasiparticle energies with relatively low computational cost. While this method does not account for self-consistency, it often yields sufficiently accurate results for many applications, and is still a significant improvement from typical *ab initio* methods. This method can be extended for more accurate results. So-called Eigenvalue-Only Self-Consistent GW (evGW) iteratively applies the G_0W_0 procedure, and reconstructs the Green's function and screened Coulomb

potential using the eigenvalues obtained from the previous calculation until self-consistency is achieved. This method improves the accuracy of the quasiparticle energies without the full computational burden of full self-consistent GW (scGW). This procedure is only more computationally demanding than G_0W_0 in that it requires multiple runs, however each of these runs has the same computational demand as the initial G_0W_0 calculation.

Improving the computational efficiency of MBPT calculations for MOFs involves a combination of strategies that optimize the use of computational resources while maintaining accuracy. By using the primitive unit cell, employing the Γ -point for flat-band insulating MOFs, applying the BG terminator, and leveraging simplified GW approximations, we can significantly reduce the computational burden of these calculations. These methods collectively enhance our ability to study complex materials like MOFs using MBPT, paving the way for more efficient and accurate investigations.

4.4 Results and Discussion

We applied a modified version of the procedure developed by Kshirsagar et al. to investigate the electronic properties of three representative flat-band MOFs: MOF-5, MIL-125, and UiO-66.⁷⁸ In order to ensure accurate structural modelling, the primitive unit cells for all three MOFs were optimized using the PBEsol functional with a 500 eV plane-wave cutoff as implemented by VASP.⁹² The structural optimizations were carried out with a $2\times 2\times 2$ k-point grid, ensuring a balance between computational cost and the accuracy of the results. Following optimization, self-consistent field (SCF) calculations and wavefunction generation were performed using Quantum ESPRESSO, employing a 40 Ry plane-wave cutoff and a $1\times 1\times 1$ k-grid. These parameters were selected to maintain computational

	E_g (G_0W_0)	E_g (evGW)	E_{opt} (BSE@ G_0W_0)	E_{opt} (Experimental)
MOF-5	7.15		4.33	4.71
MIL-125	6.09	7.24	4.06	3.88
UiO-66	5.49		3.93	4.26

Table 2. Calculated band gaps for MOF-5, MIL-125, and UiO-66. All values are given in eV.

efficiency while ensuring reliable results for the subsequent MBPT calculations, setting the stage for accurate electronic structure determination.

The G_0W_0 and BSE calculations were conducted using the YAMBO software package, leveraging its advanced capabilities for many-body calculations and computational efficiency.⁸⁹ For the G_0W_0 calculations, we included 2000 total bands (266 occupied and 1744 unoccupied) and used a BG terminator to accelerate convergence with respect to the number of bands. Additionally, the BZ integration was carried out using 3,000,000 random points for the RIM to enhance the accuracy of the Γ -only calculations. The BSE calculations, which provide insights into the optical properties of the materials, considered 96 occupied and 96 unoccupied bands to generate the Bethe-Salpeter Kernel (BSK). Only the resonant coupling between the electron-hole pairs was considered in the BSK calculation. The BSK diagonalization was referenced to the G_0W_0 -obtained band gap corrections to the GGA eigenvalues. Both the G_0W_0 and BSE calculations employed a truncated plane wave cutoff of 6 Ry to manage computational expenses, ensuring the feasibility of these demanding computations.

The G_0W_0 calculations yielded fundamental band gaps of 7.15 eV for MOF-5, 6.09 eV for MIL-125, and 5.49 eV for UiO-66 (see Table 2). As anticipated, these values were under converged due to the limitations of the G_0W_0 approximation - indeed, the calculated value for MOF-5 was nearly 1 eV lower than the value

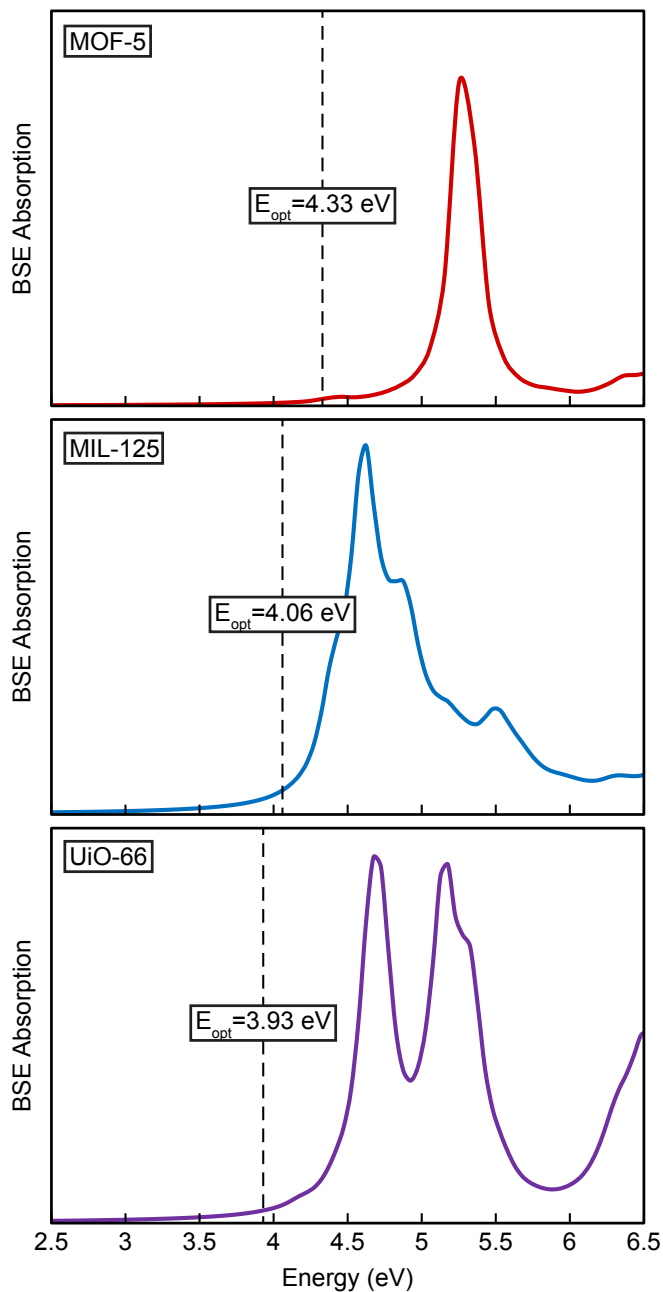


Figure 9. Simulated absorption curves for MOF-5, MIL-125, and UiO-66 calculated using the BSE. The optical band gaps (dashed lines) are determined by the lowest-energy exciton from the calculation. The gaps for MOF-5 and UiO-66 are lower than the experimental values, likely because they are underconverged with respect to the self-energy. However, the gap for MIL-125 exceeds the experimental optical gap, likely due to significant spin-orbit coupling near the conduction band minimum.

reported previously.⁷⁸ Additionally, an *evGW* calculation was performed on MIL-125 to achieve self-consistency and the calculated E_g value increased by more than 1 eV, further exacerbating the overestimation of the optical gap. The BSE calculations yielded optical band gaps of 4.33 eV for MOF-5, 4.06 eV for MIL-125, and 3.93 eV for UiO-66 (see Table 2). As these are optical band gaps (which can be measured experimentally), they are directly comparable with experimentally values. Interestingly, while the calculations underestimated the optical gaps for MOF-5 and UiO-66 as expected, the calculation overestimated the optical gap for MIL-125. This divergence from expected behavior warranted further investigation.

We hypothesize that the overestimation of the optical gap for MIL-125 can likely be attributed to spin-orbit coupling effects. MBPT calculations, particularly GW calculations, are sensitive to spin-orbit interactions due to the explicit treatment of electron exchange. In the case of MIL-125, the frontier conduction bands consist primarily of Titanium d-orbitals. Excitations populating these orbitals result in unpaired electrons in a transition metal d-orbital, which are known to experience significant spin-orbit coupling due to the motion of the high-energy electrons. This interaction likely leads to the observed overestimation of the band gap, as spin-orbit coupling can significantly affect the electronic structure of materials containing transition metals. These results underscore the need for careful consideration of computational parameters and potential physical effects such as spin-orbit coupling, especially when dealing with materials involving transition metals.

4.5 Future Directions

To further advance the understanding of the electronic properties of MOFs and improve the accuracy of MBPT calculations, several key areas warrant

exploration and refinement. One of the primary areas of interest is the role of spin-orbit coupling (SOC) in the overestimation of both E_g and E_{opt} in MIL-125. Incorporating SOC into ab initio calculations, including MBPT, poses significant computational challenges, as SOC effectively doubles the number of bands required for the calculations. This is because each band must be treated as half-filled based on Hund's rule and the Pauli Exclusion Principle. This effectively doubles the demand (at the minimum) in computational cost, further complicating the convergence process with respect to the number of bands. For this reason, the use of a BG terminator becomes essential in these calculations. Additionally, given the unique dielectric properties of MOFs it is worth exploring the application of a BG terminator not only for the self-energy but also for the dielectric screening. This approach could potentially enhance the precision of this approach for MOFs. Unfortunately, the BG terminator has shown issues in the latest version of YAMBO, and recent hardware updates have raised compilation issues with older versions.

Thorough convergence testing needs to be conducted for all investigated MOFs in the future. In this study we followed the procedure by Kshirsagar *et al.* under the assumption that they performed convergence testing on MOF-5, but convergence must be tested on each system individually and is crucial to ensure the reliability and accuracy of the results. These tests will focus on convergence with respect to the number of bands and the plane-wave energy cutoff. By systematically varying these parameters and analyzing their impact on the calculated electronic properties, we can identify the optimal settings that balance computational cost and accuracy.

Another promising avenue for future research is the investigation of quantum confinement effects in a monolayer of $\text{Ni}_3(\text{HITP})_2$, a conductive MOF that has been reported as both metallic and semiconducting in different studies. This MOF presents a unique challenge due to its indirect band gap and significant band dispersion, rendering Γ -only calculations insufficient. Furthermore, the presence of Nickel d-orbitals in the frontier bands suggests substantial spin-orbit coupling effects. Preliminary calculations using a $2 \times 2 \times 1$ k-grid and 159 unoccupied bands (300 total) yielded an E_g value that was 300 meV larger than the experimentally measured gap, which could be explained by significant spin-orbit coupling effects. Additionally, the interpolated band structure did not seem to match those previously reported in the literature, indicating the need for more comprehensive k-point sampling for accurate results.⁹³ To address the challenge of k-mesh sampling, a potential alternative solution is to use a shifted grid approach. YAMBO facilitates this by allowing the generation of a shifted grid based on an initial G_0W_0 calculation. The desired k-point is specifically targeted in this shifted grid, which then becomes the only weighted k-point. Quantum ESPRESSO can be used to generate a new set of wavefunctions from this shifted grid, followed by a subsequent G_0W_0 calculation in YAMBO using these wavefunctions and the dielectric screening obtained from the original calculation. This approach aims to enhance the accuracy of band structure calculations by incorporating additional k-points not sampled in the initial calculation.

Finally, another promising alternative for studying the excited states of MOFs involves examining fragments of MOFs as clusters and performing non-periodic calculations on them. This method can significantly reduce computational costs while preserving the accuracy of the calculations, as the electronic properties

of bulk MOFs are often maintained in MOF fragments due to the highly localized nature of their electronic states.⁷⁹ Furthermore, cluster methods allow for the application of higher levels of theory to account for relativistic effects such as spin-orbit coupling. However, the main challenge with this method is the selection of appropriate fragments from the MOF. Incorrect geometry or improper truncation of the MOF can lead to the loss of valuable electronic structure information, making it imperative to carefully choose the fragments for accurate representation of the bulk material.

Despite the inherent challenges posed by MOFs, such as their large unit cells, unique dielectric properties, and pronounced excitonic effects, our work demonstrates that MBPT can offer significant improvements over conventional methods. While the results are not yet fully conclusive, they represent a clear step forward in accurately modeling the electronic properties of MOFs. Through the strategic use of simplifying approximations, we are able to overcome some of the current limitations of MBPT. In this way we can enhance the predictive power and reliability of MBPT calculations for a wide range of MOFs, paving the way for potential industrial application.

CHAPTER V

APPLICATION: ELECTRONIC DEFECTS MEDIATED THROUGH GUEST IONS

This work was co-authored by Connor Welty, Eoghan L. Gormley, Julius J. Oppenheim, Mircea Dincă, Christopher H. Hendon, and Nicholas P. Stadie. Connor Welty collected all experimental data presented in this work, and all calculations were carried out by Eoghan L. Gormley; both of these authors contributed equally to this work. The experimental procedure used by Connor Welty was based on work previously done by Julius J. Oppenheim. Mircea Dincă, Christopher H. Hendon, and Nicholas P. Stadie are acknowledged for their advisory roles and intellectual contributions to this work.

5.1 Divergent electrically conductive pathways in 2- and 3- dimensional yttrium-based MOFs

Most metal-organic frameworks (MOFs) are electrical insulators due to energetic mismatch between the metal clusters and the ligands that support them.⁹⁴⁻⁹⁶ MOFs that do conduct electricity feature conductive pathways that permit charge mobility either through space (through π -stacking or other secondary bonding)⁹⁷ or through bond (through dative, ionic, or covalent bonds)⁹⁸, and are potentially useful as active materials in electrocatalysis⁹⁹⁻¹⁰¹ and in electrochemical energy storage devices.¹⁰²⁻¹⁰⁵ To date, the highest performing conductive MOFs feature conductivity values exceeding 100 S cm^{-1} and are through-space conductors. Effort has been invested in understanding how the spatial orientation of the MOF components dictates the crystal conductivity, however very few structure-function relationships have been revealed. A comparison of compositionally similar

frameworks with dissimilar geometries is required to elucidate basic principles that underpin electrical conduction in molecular materials.

Since electrical conductivity is primarily dictated by the number and mobility of charge carriers, interest has been focused on methods to affect band curvature and the free carrier concentration in the materials. Towards the latter, two strategies have been developed; either through electrochemical doping¹⁰⁶ or through narrowing of the electronic band gap via ligand functionalization.¹⁰⁷ With the exception of mixed-valent Fe-based frameworks, the most conductive MOFs are composed of hexasubstituted triphenylenes (e.g., hexahydroxytriphenylene, HHTP, which is referred to as HOTP once incorporated into the framework)¹⁰⁸ and their synthesis does not afford an obvious route to introduce functionalization beyond the ligating atoms themselves.^{109–111} Instead, those linkers are known to oxidize during self-assembly, introducing charge carriers through redox events. Once assembled, the charge carriers are thought to traverse the closely-spaced π -stacked layers, leading to anisotropic through-space conduction.^{93,112}

When paired with first row transition metals (e.g., Co, Ni, Cu)^{113–115}, the resultant 2D-connected layered materials exhibit among the highest known electrical conductivity of any MOFs to date. Stacking faults may disrupt the π -overlap between layers, and subsequently result in variable and reduced conductivity. Larger metal cations (e.g., La, Ho, Nd, Yb) have been shown to reduce stacking faults and form large crystals.¹¹⁶ In those cases, the metals sit between the organic sheets and, like the other 2D-conductive MOFs, those formed from *f*-block elements are known to be in-plane insulators, and out-of-plane metals. Meanwhile, some of the same metal ions can be used to form a 3D-connected cubic framework, with reported structures made from La, Eu, and Y.¹¹⁷ Thus, the

formation of iso-compositional frameworks with dissimilar connectivity, but high degrees of covalency remain key targets in MOF syntheses.

Noting that Y^{3+} supports a larger coordination sphere than its kinosymmetric azimuthal analogue (Sc^{3+}), we surmised that Y could be used to form the tightly spaced layered structure of MHOTP. In doing so, Y would be the only *d*-block element capable of reliably forming both the hexagonal and cubic structures, permitting a controlled analysis of the properties that drive electrical conduction in this class of materials. Here, we report the synthesis of YHOTP, and through comparison with its known cubic 3D counterpart, $Y_6\text{HOTP}_2$, we show that the spatial orientation of the linkers dictates the conductive pathways due to divergent linker behavior. The 2D material exhibits strong intermolecular electron and hole charge transfer coupling, while the 3D structure hosts high charge carrier concentrations.

To target the layered hexagonal 2D structure, YHOTP was obtained by mixing $Y(\text{NO}_3)_3 \cdot 6\text{H}_2\text{O}$ with HHTP in a mole ratio of 10:1 in a mixture of water, *N,N'*-dimethylimidazolidinone (DMI), and sodium acetate (NaOAc). The reaction mixture was held at 80 °C for 16 h in air, then washed and dried to obtain $(\text{YOH})_{1-x}\text{HOTP}(\text{H}_2\text{O})_n$ ($x = 0$ to 0.2; referred to as **YHOTP**). Y-vacancies are accounted for with positive values of *x*. To obtain the cubic 3D structure, YCl_3 was used in a metal:linker mole ratio of 80:1. To promote crystallization and disincentivize oxidation, the mixture was heated to 135 °C for 72 h in a nitrogen-filled glovebox in a similar solvent mixture. The resultant material, $Y_6(\text{HOTP})_2(\text{CO}_3)\text{Cl}_6$ (referred to as **$Y_6\text{HOTP}_2$**) was obtained. We surmised that the carbonate was formed during self-assembly through solvent decomposition, and trace nitrogen in the elemental analysis likely corresponds to trapped solvent within

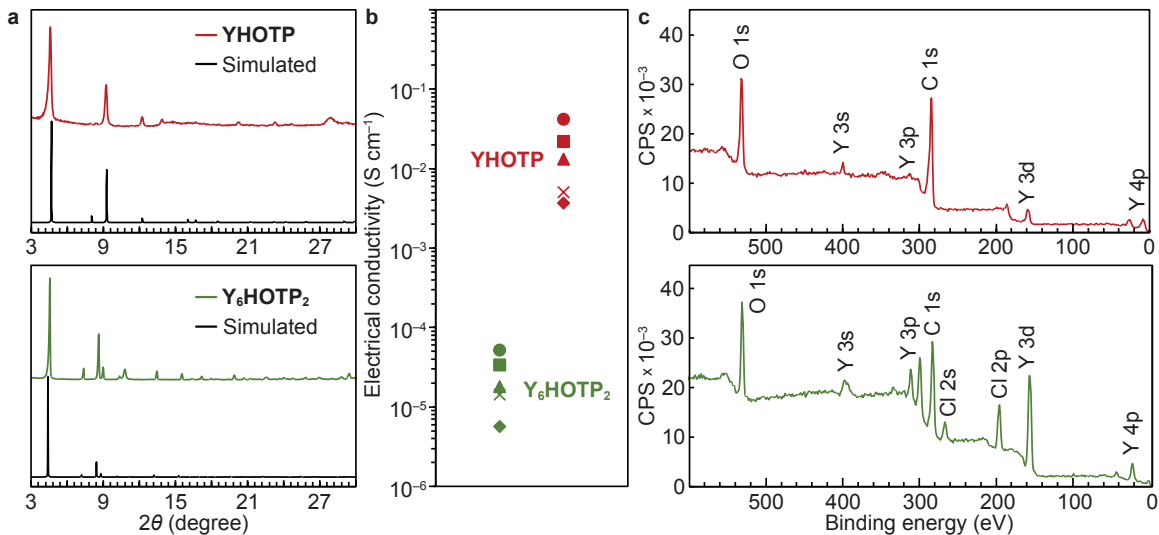


Figure 10. Materials characterization comparison of YHOTP and $Y_6\text{HOTP}_2$. (a) Powder XRD patterns reveals the microcrystalline sample is highly crystalline. (b) Pressed pellet conductivities of five different batches of each MOF. (c) XPS show the presence of Cl in the $Y_6\text{HOTP}_2$, and differences in the C 1s binding energy, attributed to differences in linker oxidation state.

the pores. Both products' chemical compositions were determined by a combination of single-crystal X-ray diffraction (XRD), X-ray photoelectron spectroscopy (XPS), and elemental analysis (Tables A.3-A.8). Microcrystalline samples of each product were found to have the same crystallographic structure as that of the bulk single crystals of each MOF (10a).

Following the reported procedure for measuring electrical conductivity on polycrystalline samples, pressed pellet conductivities were determined in pentaplicate, using a two-point probe device detailed in the Experimental Methods. The hexagonal 2D framework, YHOTP, consistently exhibited at least $1000\times$ higher conductivity than its cubic 3D counterpart, with a peak conductivity of 3×10^{-1} S cm^{-1} (10b). A similar conclusion is predicted through a comparison of the computed electronic band structures (11). However, despite the cubic material possessing an insulating ground state electronic structure¹⁰¹, the material has

surprisingly high peak conductivities of 10^{-4} S cm^{-1} , which is as much as 9 orders of magnitude greater than MOFs with comparable band gaps.⁹⁸ Thus, we surmised that another effect may be dictating the electrical conduction in these frameworks.

One explanation for this divergence could be a difference in charge carrier concentration. Since we are unable to grow single crystals large enough to perform a Hall measurement, we instead deduce the concentration from theory. By comparing the curvature of the bands in the direction with the lightest holes and electrons (Γ -A for YHOTP, Γ -X for Y_6HOTP_2), the electronic band structure calculations reveal approximately two orders of magnitude difference in effective mass of the charge carriers, with the heavier carrier belonging to Y_6HOTP_2 . Since conductivity is proportional to the charge carrier concentration and mobility, and the differences in mobility are of the same order of magnitude as the pressed-pellet conductivity, the corresponding carrier concentration must not vary by more than one order of magnitude.

Considering the composition differences, and that the XPS data presented in Figure 10c indicate that Y exists in the 3+ oxidation state, the differences in conductivity are driven by a combination of dissimilar linker oxidation state and intermolecular orientation. Since the 2D structure only forms in the presence of air, it is understood that the linker oxidizes during self-assembly. With complementary deprotonation, the hexagonal layered material hosts linkers with nominal 3- charge. This charge should result in a single unpaired electron per linker but solid-state EPR reveals YHOTP has substantially less radical character than the cubic analogue, which has been previously attributed to linker vacancies rather than native ligand-centered unpaired electrons.¹¹⁷ Thus, we surmised that the linkers

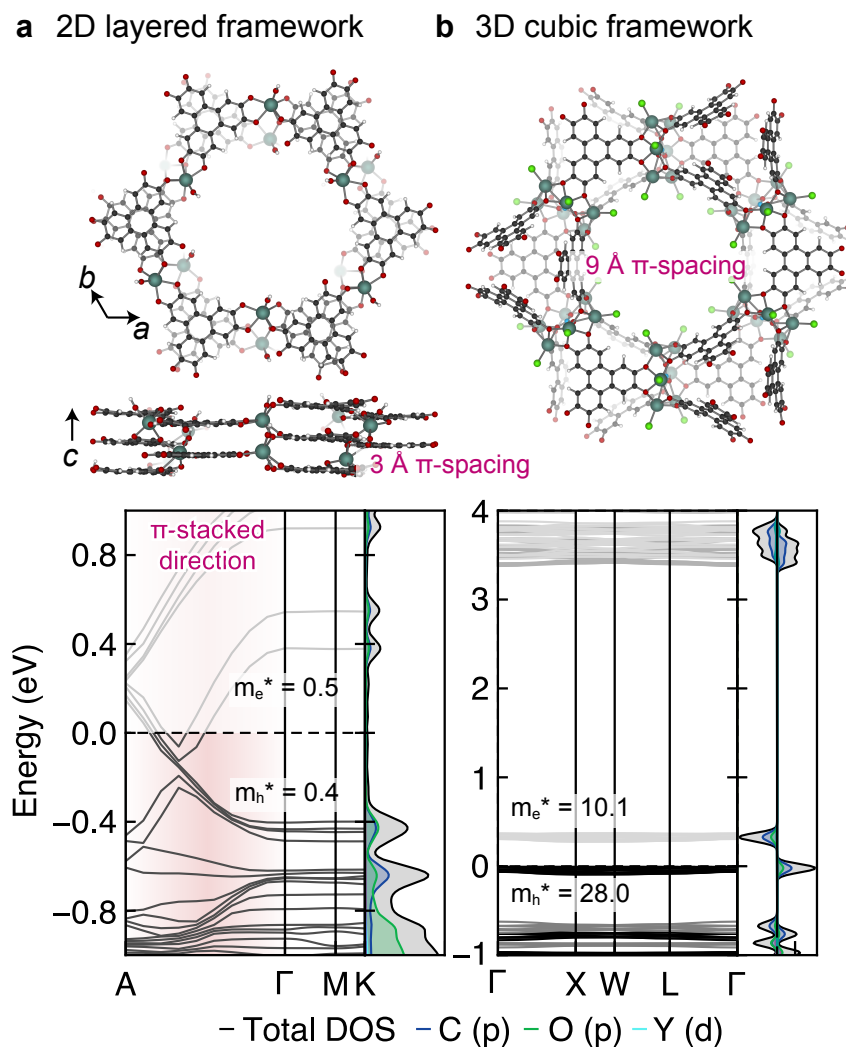


Figure 11. (a) The hexagonal (YHOTP) and (b) cubic ($Y_6\text{HOTP}_2$) crystal structures and corresponding electronic band structures. The former has pathways of through-space π -interactions, while the latter features isolated chemical motifs. The proximity of π -clouds in YHOTP is reflected in its electronic band structure, where it is predicted to be an out-of-plane metal and in-plane narrow gap insulator. $Y_6\text{HOTP}_2$ is predicted to be a narrow gap insulator due to limited spatial overlap throughout the crystal, with band edge densities localized on the ligand. Charge carrier effective masses are labeled. C, O, Cl, H, and Y, are depicted in black, red, green, white, and sage, respectively.

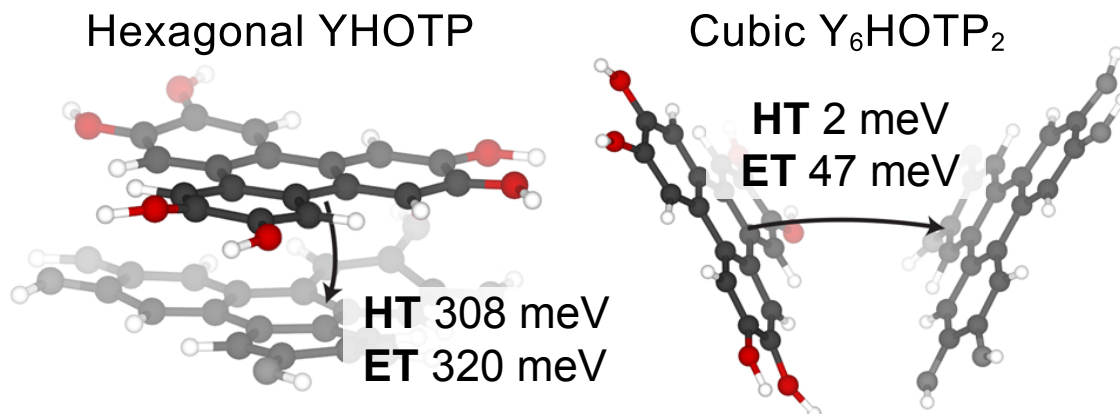


Figure 12. A comparison of the hexagonal (YHOTP) and cubic (Y_6HOTP_2) ligand charge transfer interactions. The electron coupling (ET) is consistently larger than the hole coupling (HT) driven by the propensity for the ligands to rearomatize. C, O, and H are depicted in black, red, and white, respectively.

rather exist as a delocalized 2⁻ and 4⁻ pair, an electronic configuration common in the 2D conductive family of frameworks.⁹³

Conversely, Y_6HOTP_2 forms under a nitrogen atmosphere and, hence, should be less oxidized. In its case, the linkers exhibit a formal 5⁻ charge. In this charge state, the linkers are nominally more aromatic than their oxidized counterparts and should host unpaired electrons. EPR revealed a ligand-centered radical associated with a single unpaired electron per linker. The difference concentration in mobile charge carriers cannot vary by more than one electron per linker, supporting our offering that the difference in charge carrier concentration will never exceed one order of magnitude. This may be the origin of elevated conductivity in Y_6HOTP_2 compared to other similarly flat-banded materials — the linker is likely installing large numbers of charge carriers.

Both YHOTP and Y_6HOTP_2 are through-space conductors but have different ligand geometries. This is generally interesting because intermolecular distances usually increase with increasing porosity, and these frameworks serve

as a platform to understand how orientation impacts through-space conduction, and whether having spatially separated ligands can permit charge transfer between them. To do so, we turn to Marcus theory to assess the charge transfer coupling between neighboring ligands. We created two models from the bulk crystals, each featuring a pair of linkers with their geometry remaining fixed, as imposed by the geometry obtained for the crystal.¹¹⁸ The linkers were then protonated to create a charge-neutral system (Figure 12). The electron and hole couplings for each cluster were calculated using the open-source code CATNIP. For the cubic system, the hole transfer coupling was found to be 2 meV, while the electron coupling was 47 meV. Both of these values are extremely weak. The hexagonal crystal featured coupling energies of 320 and 308 meV for the electron and hole, respectively. These values are relatively large and are comparable to other organic charge transfer aggregates.^{119,120} Together we surmise that for the hexagonal system, the mobility of the charge carriers is high reflected in the band structure and also in the hole and electron coupling energies. Conversely, the cubic system features weak couplings, and highly localized electronic bands.

By further computing the reorganization energies for both the electron and hole transfer to the protonated ligand, we are able to then compute the rate of charge transfer. Since the linkers are the same for both the cubic and hexagonal phase, the reorganization energies should be roughly the same for both systems. The electron and hole reorganization energies were found to be 587 meV and 491 meV, respectively. From Marcus theory, the Arrhenius rates for the hexagonal system were $7.47 \times 10^{12} \text{ s}^{-1}$ and $1.92 \times 10^{13} \text{ s}^{-1}$ for electron and hole transfer, respectively. The cubic material exhibits rates 2-4 orders of magnitude slower, $1.63 \times 10^{11} \text{ s}^{-1}$ and $6.57 \times 10^8 \text{ s}^{-1}$ for electron and hole transfer, respectively. These

data suggest differences in conductivity are further dictated by differences in charge transfer rates, associated with differences in geometry.

Thus, through the isolation of a Y-based hexagonal layered conductor, YHOTP, and by comparison to its cubic counterpart, $Y_6\text{HOTP}_2$, we can make several conclusions. Given that the coupling energies in the latter are extremely weak, but its charge carrier effective masses are comparable to other flat-band insulators, the reason for its surprisingly high conductivity ($10^{-4} \text{ S cm}^{-1}$) is likely due to the differences in linker oxidation state, which we believe is directly influencing the charge carrier concentration in the MOF. Further, given that the linker is more aromatic in the 3D framework — a property that is typically inversely proportional to conductivity¹²¹ - this strongly indicates that linker redox is a critical and primary deciding factor dictating electrical charge conduction in MOFs, and highlights that the interplay between linker redox potential and orientation dictates bulk conductivity. Looking forward, these findings help rationalize why the HOTP-based materials are the best conductors in their class and should instruct future efforts to diversify ligand design.

5.2 Methods Section

5.2.1 Synthesis of YHOTP and $Y_6\text{HOTP}_2$. Prior to synthesis, HHTTP was dissolved in DMI and recrystallized from hot H₂O to remove impurities. For the synthesis of YHOTP, 0.63 mmol of $Y(\text{NO}_3)_3 \cdot 6\text{H}_2\text{O}$ was dissolved in 3 mL of deionized H₂O in a 20 mL scintillation vial in air. 0.063 mmol of HHTTP was dissolved in 0.8 mL of DMI and the two solutions were combined. A solution of 0.189 mmol of NaOAc dissolved in 0.6 mL of deionized H₂O was added to the mixture. The mixture was transferred to a 15 mL pressure tube containing sanded glass slides and heated at 80 °C for 16 h. For the synthesis of $Y(\text{NO}_3)_3 \cdot 6\text{H}_2\text{O}$,

4.8 mmol of $\text{YCl}_3 \cdot 6\text{H}_2\text{O}$ was dissolved in 3 mL of deionized H_2O and added to a solution of 0.06 mmol of HHTP dissolved in 0.8 mL of DMI in a N_2 -filled glovebox. 0.18 mmol of NaOAc dissolved in 0.6 mL of deionized H_2O was added to the mixture and then the solution was filtered into a 15 mL pressure tube containing sanded glass slides using a $0.2 \mu\text{m}$ PTFE filter. The reaction was heated at $135 \text{ }^\circ\text{C}$ for 72 h and then brought back into the N_2 -filled glovebox. The mother liquor was decanted and the resulting powder was washed $3\times$ with deoxygenated DI water and $3\times$ with deoxygenated methanol. Both samples were activated by drying under vacuum at $90 \text{ }^\circ\text{C}$ for 1 h.

5.2.2 Electrical Conductivity Measurements. Two-point probe measurements were made on activated pressed pellets using a custom apparatus previously described.¹²² 2-3 mg of powder was loaded into a glass tube and compressed by copper rods that also served as leads. The apparatus was compressed using a hydraulic press while dual linear I-V curves were obtained using a Keithley 2450 source meter. The pellets were further compressed between I-V measurements until the change in resistance was low. The conductivity measurements were made in pentaplicate.

5.2.3 Materials Characterization. Powder XRD measurements were performed using a Bruker D8 Advance diffractometer with $\text{Cu K}\alpha_{1,2}$ radiation ($\lambda = 1.54 \text{ \AA}$) in reflection geometry. EPR was performed using a Magnettech ESR5000 spectrometer with a microwave power of 20 mW and an amplitude of 0.3 G.

5.2.4 Computational Methods. Density functional theory (DFT) calculations were performed using the Vienna ab initio Simulation Package (VASP).⁹² All structures were optimized using the PBEsol functional with a plane

wave energy cutoff of 500 eV, and an ionic convergence criterion of -0.005 eV. Bulk YHOTP was optimized using a $2 \times 2 \times 5$ k-mesh to sample the first Brillouin zone, and all structures involving Y_6HOTP_2 used a Γ -only k-mesh due to the size of the unit cell and the resulting computational cost of the calculation. Y_6HOTP_2 is charge-balanced by disorder guest molecules in the pores. Here, we achieved charge balance by adding one Cl per Y, capping the open metal sites in the metal clusters. The electron count was adjusted by the removal of 4 electrons, achieving charge neutrality. This strategy significantly reduced the net charge of the structure in the calculations, aligned with the experimental number of charges on each linker from EPR, and is justified due to the presence of Cl in the elemental analysis. We also tested an alternate charge balancing method, by removing the chlorides and subtracting 20 electrons (2.5 per linker) from the system. The former strategy yielded a plausible structure that was then used for the electronic band structure calculation. The HSEsol hybrid functional¹²³ was used to compute the band structures. Charge transfer calculations were performed on clusters extracted from the geometrically equilibrated crystal structures. Initial calculations were performed using the ZINDO method, as implemented in Gaussian09,¹²⁴ and yielded similar coupling energies as computed using the B3LYP functional and 6-31G basis set (Figure 12). Coupling energies were calculated using the Molecular Orbital Overlap (MOO) method.

CHAPTER VI

CONCLUSION

Modifying conventional computational methods is crucial for better understanding defects in molecular materials such as MOFs, as these systems present unique challenges. Due to the intrinsic complexity of these materials, they can accrue a wide and exotic range of defects ranging from conventional interstitial atoms and atom-substitutions to entire linker vacancies or guest ions and guest molecules. Additionally, the chemical reactions involved in forming these defects often involve the cleaving and formation of covalent bonds. As a result of these complexities, conventional reference state methods for the chemical potential term in Equation 1.1 become inconsistent and potentially inaccurate, therefore necessitating more careful treatment of the chemistry involved in the formation processes for these defects. The shortcomings of conventional *ab initio* methods are also compounded by many molecular materials due to strong excitonic effects and highly-localized electronic states. Attempts to reconcile these difficulties through the use of higher levels of theory are often further complicated by the large sizes of these systems, and therefore simplifying approximations and clever computational strategies are essential to the success of these attempts.

Several modifications to computational methods were implemented to account for the unique challenges presented by molecular materials. Existing techniques for calculating the chemical potentials for defects were adapted to better model the chemical reactions involved in defect formation, and higher levels of theory were implemented to more accurately model the unique electronic structure of MOFs. These adjustments were necessary to handle the large unit cells, complex chemical environments, and intricate structural dynamics that molecular materials

often exhibit. These refinements allowed for a more realistic representation of the materials' behavior that better aligns with experimental conditions.

The treatment of chemical potential was approached in an intuitive way that implicitly accounted for the cleavage and formation of covalent bonds, thereby including the thermodynamic landscape of the reaction directly into the formation energy of the defect. Conventional methods can fall short in this regard due to the convention of treating all atoms of a given type as chemically equivalent in the defect reservoir. Additionally, referencing chemical potentials to bulk materials implicitly considers reactions that are not present in experimental settings, which requires careful consideration of the stoichiometry of the system to ensure cancellation with Hess's law.

Electronic properties and band gaps in molecular materials present another layer of complexity. Conventional *ab initio* methods frequently underestimate fundamental band gaps, especially in MOFs, due to strong excitonic effects and highly localized electronic states. To address this, higher levels of theory such as MBPT must be employed. MBPT offers a more accurate representation of electronic properties, enabling better predictions of defect behavior, but with a much steeper computational cost. The necessity of using such advanced theoretical approaches has been underscored by the significant discrepancies observed with simpler methods in previous studies on hybrid organic-inorganic perovskites. This research not only highlighted the limitations of conventional methods but also provided a pathway to more accurate and reliable predictions through the integration of MBPT with several approximations and simplifications to reduce computational cost.

One of the most common uses for defects in materials is to modify charge carrier concentrations to alter their conductivities. A case study on Yttrium-based MOFs with HHTP linkers highlighted the practical implications of these modifications. Different conductivities were observed in the 2D and 3D phases, with the 3D phase's conductivity being primarily influenced by charge carriers introduced by guest ions in the pores - a unique type of defect behavior primarily observed in porous materials. This example illustrated the significant impact that defects and external factors can have on material properties, reinforcing the importance of accurate defect modeling in MOFs. The study provided insights into how the presence of defects and guest ions can alter electronic properties such as conductivity, demonstrating the practical applications of computational methods. While this case study primarily analyzed the conductivity through the use of fragment clusters of MOF linkers, it underscores the broader applicability of the computational techniques to other MOFs and molecular materials.

Theoretical implications of these findings extend to a deeper understanding of defects in molecular materials. This research seeks to enhance the ability of computational methods to predict material behavior by improving the accuracy of defect modeling, paving the way for more reliable studies of MOFs and similar materials. This work sets a precedent for further development of computational techniques tailored to the unique properties of molecular materials. Practical applications of these improved predictions are far-reaching in material science and industry. Insights gained from accurate modeling of defect properties can inform the design and optimization of MOFs for various applications, including catalysis, gas storage, and electronic devices. The ability to develop materials with tailored functionalities is crucial for technological advancements and industrial processes.

These practical implications highlight the value of the developed computational methods for both academic research and industrial practice.

Despite the significant advancements made, certain limitations are present. The procedure for calculating the chemical potential still requires experimental verification, and in terms of application requires that the full reactions (including product *and* reactant species) for forming defects must be known. Further exploration of a broader range of materials and defect types is essential to fully validate and expand upon these findings. Additionally, reaction conditions such as temperature and pressure were not factored into these calculations, meaning that entropy might play a significant role. MBPT calculations carried out here rely heavily on certain approximations, and may still be limited by system size. Larger MOFs may still prove to be difficult in this regard, and MOF systems with dispersive bands require denser k-point sampling, which can add tremendously to computational cost. The scalability of the modified methods for even larger and more complex systems remains a challenge that future research should aim to overcome. Identifying and addressing these limitations will enhance the robustness and applicability of the computational approaches presented here.

Future research should focus on refining computational methods and exploring their applicability to a wider range of molecular materials. Developing more sophisticated models to handle the complexity of different MOFs and their defects is crucial. Experimental collaborations are also essential for validating computational predictions and providing empirical data to support theoretical findings. Combining computational and experimental approaches will achieve a more comprehensive understanding of defect behavior in molecular materials. Future work should also explore new computational strategies and technologies

to further improve the efficiency and accuracy of defect modeling, ensuring that the methods remain at the forefront of the field. Advancing computational techniques to meet the unique demands of molecular materials is essential for future discoveries and technological innovations, and these findings have the potential to offer valuable insights into the design and optimization of MOFs. This dissertation includes previously published coauthored material.

APPENDIX

SUPPLEMENTARY INFORMATION FOR CHAPTER V

	YHOTP	Y ₆ HOTP ₂
Chemical Formula	Y _{1.17} (C ₁₈ H ₆ O ₆)(H ₂ O) _{1.17}	Y ₆ (CO ₃)(C ₁₈ H ₆ O ₆) ₂ (H ₂ O) ₆ Cl ₆
Node SBU	YO ₆ (H ₂ O) = [Y] ³⁺	Y ₆ O ₁₂ (CO ₃)(H ₂ O) ₆ (Cl) ₆ = [Y ₆] ¹⁰⁺
Linker SBU	C ₁₈ H ₆ O ₆ = [HOTP] ³⁻	C ₁₈ H ₆ O ₆ = [HOTP] ⁵⁻
Chemical Composition	Y _{1.17} C ₁₈ H _{8.34} O _{7.17}	Y ₆ C ₃₇ H ₁₂ O ₂₁ Cl ₆
Crystal System	Hexagonal	Cubic
Space Group	P6cc (no. 184)	Fd-3m (no. 227)
<i>a</i> (Å)	21.901	34.548
<i>b</i> (Å)	21.901	34.548
<i>c</i> (Å)	6.1825	34.548
α (°)	90	90
β (°)	90	90
γ (°)	120	90
Density (mL g ⁻¹)	1.15	0.99
Network Connectivity	3D	3D
Pore Connectivity	1D	3D

Table A.3. Structural models for YHOTP and Y₆HOTP₂. Coordinated water molecules (identified as one-sided O atoms in single-crystal studies) are included. Charge balance is indicated based on EPR results.

This is a sample citation:[?] .

	Y ₆ HOTP ₂
Empirical Formula	Y ₆ C ₃₇ H ₁₂ O ₂₁ Cl ₆
Formula Weight	1538.64
Temperature (K)	100
Crystal System	Cubic
Space Group	Fd-3m
<i>a</i> (Å)	34.548(6)
<i>b</i> (Å)	34.548(6)
<i>c</i> (Å)	34.548(6)
α (°)	90
β (°)	90
γ (°)	90
<i>V</i> (Å ³)	41235(2)
<i>Z</i>	16
Density (g mL ⁻¹)	0.9914
F(000)	126888
μ (mm ⁻¹)	7.475
Radiation	Cu K α ($\lambda = 1.54184$ Å)
2θ range (°)	4.43–158.724
Reflections collected	151190

Table A.4. Single-crystal XRD structure refinement of Y₆HOTP₂.

	Y	C	N	O	Cl
Y ₆ HOTP ₂	5.1	57.2	6.9	19.1	11.7
YHOTP	7.1	69.5	2.9	20.5	0.0

Table A.5. Chemical composition of YHOTP and Y₆HOTP₂ (in at%, excluding H) determined by XPS analysis.

	Y	C	N	O	Cl
Y ₆ HOTP ₂	7.5	32.1	4.5	50.9	5.0
YHOTP	3.1	44.8	1.1	50.3	0.7

Table A.6. Chemical composition of YHOTP and Y₆HOTP₂ (in at%, excluding H) determined by combustion analysis (C, N), inductively-coupled plasma (ICP) analysis (Y, Cl), and the difference (O).

	Y	C	N	O	Cl
$Y_6\text{HOTP}_2$	8.6	52.9	0.0	30.0	8.6
YHOTP	4.4	68.3	0.0	27.2	0.0

Table A.7. Model chemical composition of YHOTP and $Y_6\text{HOTP}_2$ (in at%, excluding H) based on the combined experimental analyses: XPS, CHN, ICP-OES, and XRD.

	Y	C	N	O	Cl
$YCl_3 \cdot 6H_2O$	8.1	39.1	0.0	32.0	20.8
YHOTP	4.4	68.3	0.0	27.2	0.0

Table A.8. Chemical composition of $YCl_3 \cdot 6H_2O$ (in at%) determined by energy dispersive X-ray spectroscopy (EDX) using a 10 kV accelerating voltage.

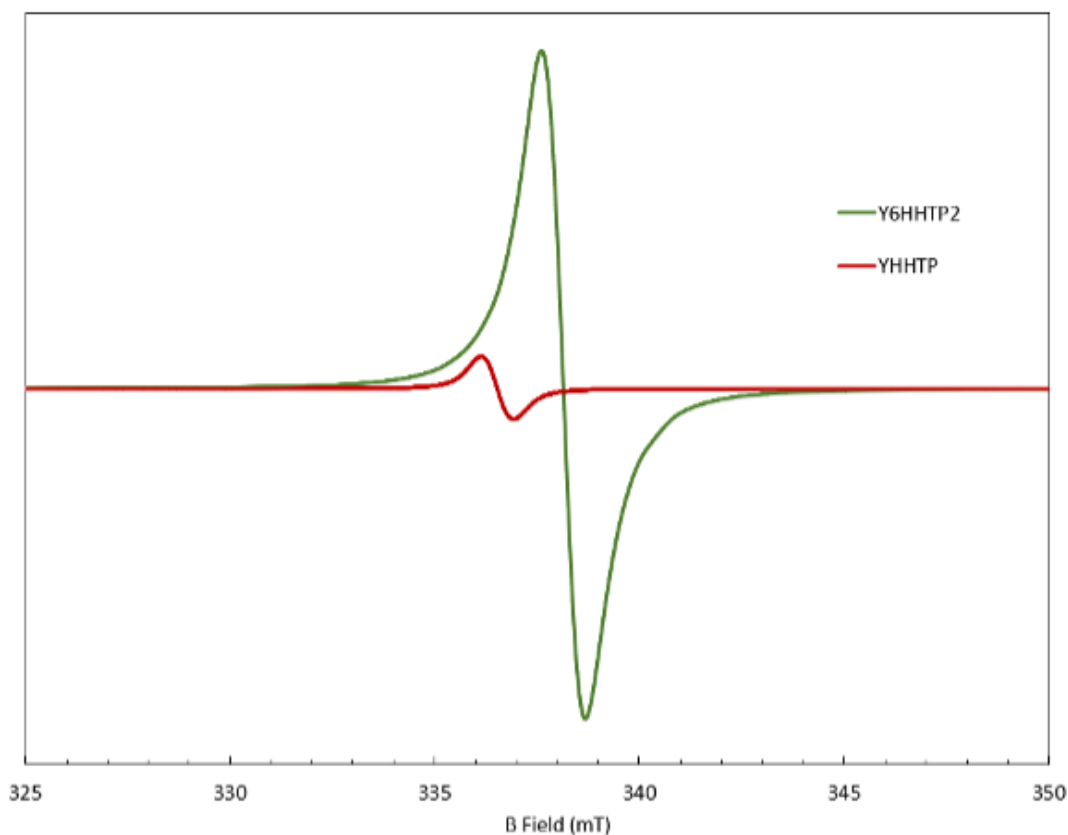


Figure A.13. EPR spectra of YHOTP and $Y_6\text{HOTP}_2$ between 325-350 mT, normalized by mass.

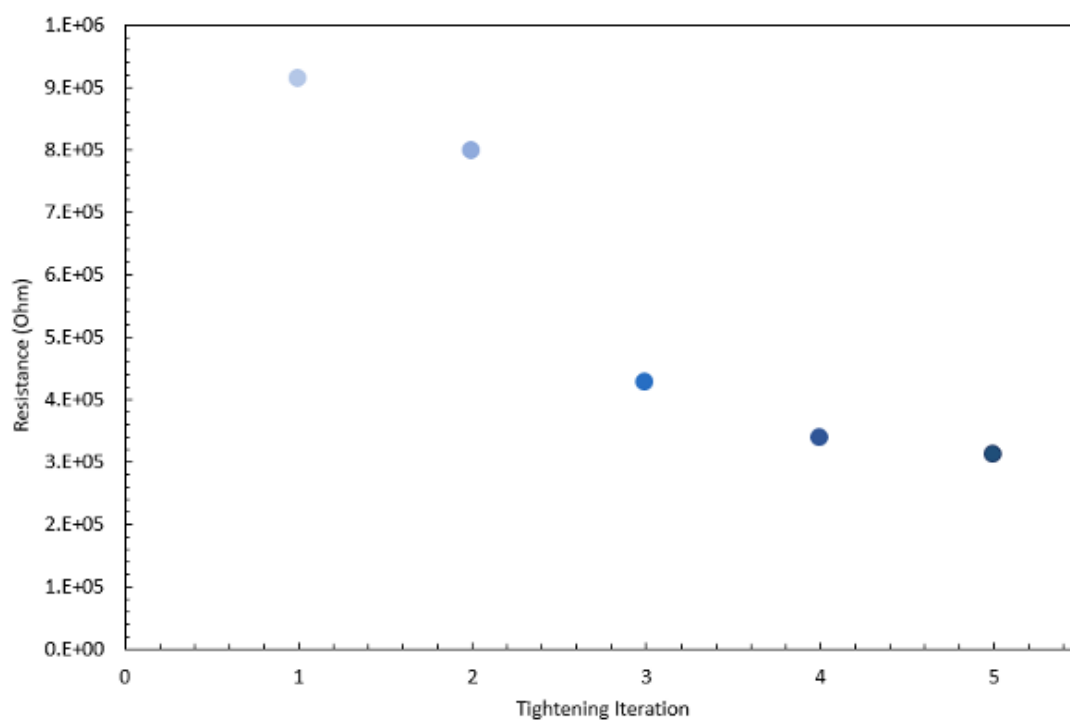


Figure A.14. Iterative two-point probe measurements of packed pellet resistance of YHOTP (I-V curve swept from -1.0 to 1.0 V), upon hand tightening between measurements.

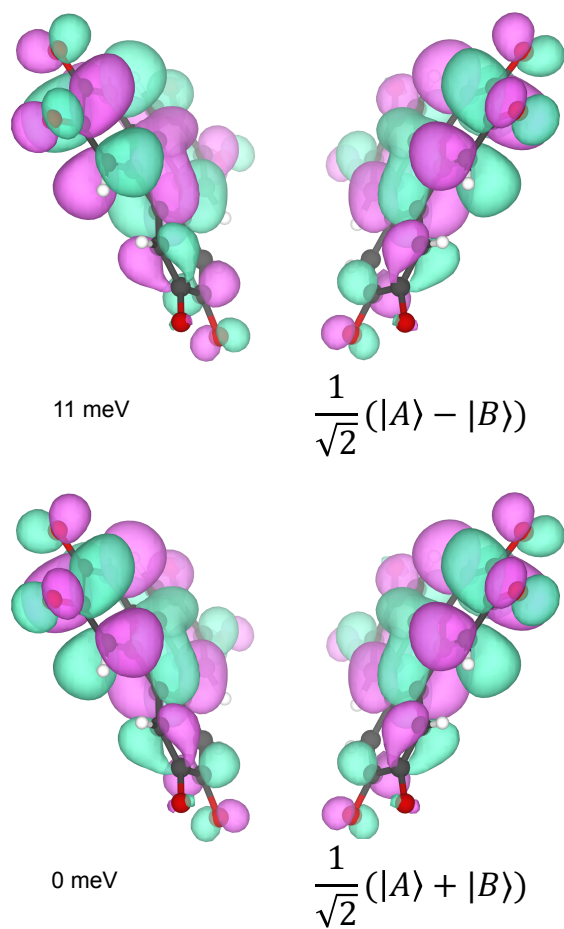


Figure A.15. Molecular orbitals of HOTP placed at positions consistent with Y6HOTP2, consisting of symmetric (bottom) and antisymmetric (top) combinations of the individual linker molecular orbitals. The antisymmetric state is higher in energy, suggesting J-like aggregation behavior.

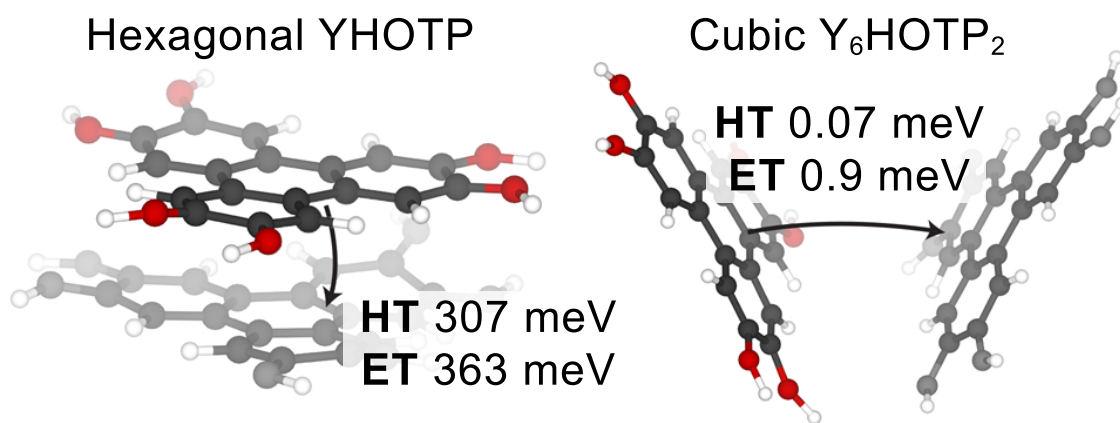


Figure A.16. A comparison of the hexagonal (YHOTP) and cubic (Y6HOTP2) ligand charge transfer interactions calculated using the ZINDO method. The electron and hole couplings were calculated using the open-source code py-MOO. The HT values are comparable to Figure 3, but this method calculated smaller ET values. Despite this, the ET values are still consistently larger than the HT values.

REFERENCES CITED

- (1) Freysoldt, C.; Grabowski, B.; Hickel, T.; Neugebauer, J.; Kresse, G.; Janotti, A.; Van de Walle, C. G. *Rev. Mod. Phys.* **2014**, *86*, 253–305.
- (2) Dragoni, D.; Daff, T. D.; Csányi, G.; Marzari, N. *Phys. Rev. Materials* **2018**, *2*, 013808.
- (3) Jones, R. O. *Reviews of Modern Physics* **2014**, *86*, 253–287.
- (4) Prasad, M. K.; Al-Ani, O. A.; Goss, J. P.; Mar, J. D. *Phys. Rev. Materials* **2023**, *7*, 094003.
- (5) Zhang, W.; Liu, Y.; Chen, H. *Computational Materials Science* **2023**, *210*, 111289.
- (6) Kim, S.; Hood, S. N.; Park, J.-S.; Whalley, L. D.; Walsh, A. *Journal of Physics: Energy* **2020**, *2*, 036001.
- (7) Goyal, A.; Gorai, P.; Peng, H.; Lany, S.; Stevanović, V. *Computational Materials Science* **2017**, *130*, 1–9.
- (8) Liang, Y.; Cui, X.; Li, F.; Stampfl, C.; Ringer, S. P.; Huang, J.; Zheng, R. *Journal of the American Chemical Society* **2022**, *144*, 6770–6778, PMID: 35385287.
- (9) Jiang, Y.; Lee, S. B.; Herzog-Arbeitman, J.; Yu, J.; Feng, X.; Hu, H.; Călugăru, D.; Brodale, P. S.; Gormley, E. L.; Vergniory, M. G.; Felser, C.; Blanco-Canosa, S.; Hendon, C. H.; Schoop, L. M.; Bernevig, B. A. *Phys. Rev. B* **2023**, *108*, 235127.
- (10) Fu, Y.; Yao, Y.; Forse, A. C.; Li, J.; Mochizuki, K.; Long, J. R.; Reimer, J. A.; De Paëpe, G.; Kong, X. *Nature Communications* **2023**, *14*, 2386.
- (11) Van de Walle, C. G. *Phys. Rev. Lett.* **2000**, *85*, 1012–1015.
- (12) Broberg, D.; Bystrom, K.; Srivastava, S.; Dahlliah, D.; Williamson, B. A. D.; Weston, L.; Scanlon, D. O.; Rignanese, G.-M.; Dwaraknath, S.; Varley, J.; Persson, K. A.; Asta, M.; Hautier, G. *npj Computational Materials* **2023**, *9*, 72.
- (13) Fabrizio, K.; Le, K. N.; Andreeva, A. B.; Hendon, C. H.; Brozek, C. K. *ACS Materials Letters* **2022**, *4*, 457–463.
- (14) Jain, V. K. *Solid State Physics*; Springer: Cham, 2022; p Chapter 10.

- (15) Liu, Q.; Dalpian, G. M.; Zunger, A. *Physical Review Letters* **2019**, *122*, 106403.
- (16) Davydov, S. Y.; Obraztsov, P. A.; Kochetov, S. N. *Physical Review B* **2015**, *91*, 054301.
- (17) Karsten, K.; Ehrhart, P. *Physical Review B* **1995**, *51*, 10508.
- (18) Walsh, A. *npj Computational Materials* **2021**, *7*, 72.
- (19) Leslie, M.; Gillan, N. J. *Journal of Physics C: Solid State Physics* **1985**, *18*, 973–982.
- (20) Makov, G.; Payne, M. C. *Phys. Rev. B* **1995**, *51*, 4014–4022.
- (21) Freysoldt, C.; Neugebauer, J.; Van De Walle, C. G. *Physical Review Letters* **2009**, *102*, 016402.
- (22) Kumagai, Y.; Oba, F. *Physical Review B* **2014**, *89*, 195205.
- (23) Suo, Z.-J.; Luo, J.-W.; Li, S.-S.; Wang, L.-W. *Physical Review B* **2020**, *102*, 174110.
- (24) Chagas Da Silva, M.; Lorke, M.; Aradi, B.; Farzalipour Tabriz, M.; Frauenheim, T.; Rubio, A.; Rocca, D.; Deák, P. *Physical Review Letters* **2021**, *126*, 076401.
- (25) De Vita, A.; Gillan, M. J.; Lin, J. S.; Payne, M. C.; Štich, I.; Clarke, L. J. *Physical Review B* **1992**, *46*, 12964–12973.
- (26) Lento, J.; Pesola, M.; Mozos, J.-L.; Nieminen, R. M. *Applied Physics Letters* **2000**, *77*, 232–234.
- (27) Hoang, K. *Journal of Applied Physics* **2022**, *131*, 015705.
- (28) Srivastava, S.; Uberuaga, B. P.; Asta, M. *The Journal of Physical Chemistry C* **2023**, *127*, 17460–17472.
- (29) Li, G.; Liu, Z.-Y.; Zhang, C.-X.; Cai, X.; Yan, L.; Zhang, C.; Deng, H.-X. *Journal of Applied Physics* **2023**, *133*, 125701.
- (30) Lenzi, V.; Silva, J. P. B.; Šmíd, B.; Matolín, V.; Istrate, C. M.; Ghica, C.; MacManus-Driscoll, J. L.; Marques, L. *ENERGY & ENVIRONMENTAL MATERIALS* **2024**, *7*, e12500.
- (31) Bellman, R. E. *Dynamic Programming*; Princeton University Press: Princeton, NJ, 1957.
- (32) Cioslowski, J., Ed. *Many-Electron Densities and Reduced Density Matrices*; Mathematical and Computational Chemistry; Springer New York, NY, 2000.

- (33) Ma, Q.; Werner, H.-J. *WIREs Computational Molecular Science* **2018**, *8*, e1371.
- (34) Helgaker, T.; Jørgensen, P.; Olsen, J. *Molecular Electronic-Structure Theory*; Wiley: Chichester, UK, 2000.
- (35) Szabo, A.; Ostlund, N. S. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Dover Publications: Mineola, NY, 2012.
- (36) Shavitt, I.; Bartlett, R. J. *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*; Cambridge University Press: Cambridge, UK, 2009.
- (37) Levine, I. N. *Quantum Chemistry*, 7th ed.; Pearson: Boston, MA, 2013.
- (38) McWeeny, R. *Methods of Molecular Quantum Mechanics*, 2nd ed.; Academic Press: San Diego, CA, 1992.
- (39) Hohenberg, P.; Kohn, W. *Physical Review* **1964**, *136*, B864–B871.
- (40) Kohn, W.; Sham, L. J. *Physical Review* **1965**, *140*, A1133–A1138.
- (41) Parr, R. G.; Yang, W. *Density-Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1994.
- (42) Perdew, J. P.; Zunger, A. *Physical Review B* **1981**, *23*, 5048–5079.
- (43) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Physical Review Letters* **1996**, *77*, 3865–3868.
- (44) Burke, K. *The Journal of Chemical Physics* **2012**, *136*, 150901.
- (45) Becke, A. D. *The Journal of Chemical Physics* **1993**, *98*, 5648–5652.
- (46) Becke, A. D. *The Journal of Chemical Physics* **1993**, *98*, 1372–1377.
- (47) Lee, C.; Yang, W.; Parr, R. G. *Physical Review B* **1988**, *37*, 785–789.
- (48) Perdew, J. P.; Ernzerhof, M.; Burke, K. *The Journal of Chemical Physics* **1996**, *105*, 9982–9985.
- (49) Payne, M. C.; Teter, M. P.; Allan, D. C.; Arias, T. A.; Joannopoulos, J. D. *Reviews of Modern Physics* **1992**, *64*, 1045–1097.
- (50) Martin, R. M. *Electronic Structure: Basic Theory and Practical Methods*; Cambridge University Press: Cambridge, UK, 2004.
- (51) Troullier, N.; Martins, J. L. *Physical Review B* **1991**, *43*, 1993–2006.

- (52) Goedecker, S. *Physical Review B* **1996**, *54*, 1703–1710.
- (53) Giannozzi, P.; Baroni, S.; Bonini, N.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Chiarotti, G. L.; Cococcioni, M.; Dabo, I.; others *Journal of Physics: Condensed Matter* **2009**, *21*, 395502.
- (54) Giannozzi, P.; Andreussi, O.; Brumme, T.; Bunau, O.; Buongiorno Nardelli, M.; Calandra, M.; Car, R.; Cavazzoni, C.; Ceresoli, D.; Cococcioni, M.; others *Journal of Physics: Condensed Matter* **2017**, *29*, 465901.
- (55) Martin, R. M.; Reining, L.; Ceperley, D. M. *Interacting Electrons: Theory and Computational Approaches*; Cambridge University Press, 2016.
- (56) Hedin, L. *Physical Review* **1965**, *139*, A796.
- (57) Onida, G.; Reining, L.; Rubio, A. *Reviews of Modern Physics* **2002**, *74*, 601.
- (58) Weber, J. R.; Koehl, W. F.; Varley, J. B.; Janotti, A.; Buckley, B. B.; de Walle, C. G. V.; Awschalom, D. D. *Proceedings of the National Academy of Sciences* **2010**, *107*, 8513–8518.
- (59) Tuller, H. L.; Bishop, S. R. *Annual Review of Materials Research* **2011**, *41*, 369–398.
- (60) Park, J. S.; Kim, S.; Xie, Z.; Walsh, A. *Nature Reviews Materials* **2018**, *3*, 194–210.
- (61) Mosquera-Lois, I.; Kavanagh, S. R.; Walsh, A.; Scanlon, D. O. *npj Computational Materials* **2023**, *9*, 25.
- (62) Alkauskas, A.; McCluskey, M. D.; Van de Walle, C. G. *Journal of Applied Physics* **2016**, *119*, 181101.
- (63) Han, D.; Zhu, B.; Cai, Z.; Spooner, K. B.; Rudel, S. S.; Schnick, W.; Bein, T.; Scanlon, D. O.; Ebert, H. *Matter* **2024**, *7*, 158–174.
- (64) Dahliah, D.; Brunin, G.; George, J.; Ha, V.-A.; Rignanese, G.-M.; Hautier, G. *Energy & Environmental Science* **2021**, *14*, 5057–5073.
- (65) Xiong, Y.; Bourgois, C.; Sheremetyeva, N.; Chen, W.; Dahliah, D.; Song, H.; Zheng, J.; Griffin, S. M.; Sipahigil, A.; Hautier, G. *Science Advances* **2023**, *9*, eadh8617.
- (66) Kumagai, Y.; Tsunoda, N.; Takahashi, A.; Oba, F. *Phys. Rev. Mater.* **2021**, *5*, 123803.

- (67) Sluydts, M.; Pieters, M.; Vanhellefont, J.; Van Speybroeck, V.; Cottenier, S. *Chemistry of Materials* **2017**, *29*, 975–984, Publisher: American Chemical Society.
- (68) Zhang, J.; Zhang, X.; Turiansky, M. E.; Van de Walle, C. G. *PRX Energy* **2023**, *2*, 013008.
- (69) Buckeridge, J.; Scanlon, D.; Walsh, A.; Catlow, C. *Computer Physics Communications* **2014**, *185*, 330–338.
- (70) Willis, J.; Spooner, K. B.; Scanlon, D. O. *Applied Physics Letters* **2023**, *123*.
- (71) Boese, A. D.; Boese, R. *Crystal Growth & Design* **2015**, *15*, 1073–1081.
- (72) Sanderson, R. *Chemical bonds and bonds energy*; Elsevier, 2012; Vol. 21.
- (73) Bordwell, F. G.; Cheng, J. P.; Satish, A. V.; Twyman, C. L. *The Journal of Organic Chemistry* **1992**, *57*, 6542–6546.
- (74) Dan-Hardi, M.; Serre, C.; Frot, T.; Rozes, L.; Maurin, G.; Sanchez, C.; Férey, G. *Journal of the American Chemical Society* **2009**, *131*, 10857–10859, PMID: 19621926.
- (75) Mancuso, J. L.; Fabrizio, K.; Brozek, C. K.; Hendon, C. H. *Chem. Sci.* **2021**, *12*, 11779–11785.
- (76) Ketterson, J. B. *The Physics of Solids*; Oxford University Press, 2016; Chapter 40.2, pp 798–799.
- (77) Mancuso, J. L.; Mroz, A. M.; Le, K. N.; Hendon, C. H. *Chemical Reviews* **2020**, *120*, 8641–8715, PMID: 32672939.
- (78) Kshirsagar, A. R.; Blase, X.; Attacalite, C.; Poloni, R. *The Journal of Physical Chemistry Letters* **2021**, *12*, 4045–4051, PMID: 33881873.
- (79) Ingham, M.; Aziz, A.; Di Tommaso, D.; Crespo-Otero, R. *Mater. Adv.* **2023**, *4*, 5388–5419.
- (80) Filip, M. R.; Giustino, F. *Phys. Rev. B* **2014**, *90*, 245145.
- (81) Leppert, L. *The Journal of Chemical Physics* **2024**, *160*, 050902.
- (82) Davies, C. L.; Filip, M. R.; Patel, J. B.; Crothers, T. W.; Verdi, C.; Wright, A. D.; Milot, R. L.; Giustino, F.; Johnston, M. B.; Herz, L. M. *Nature Communications* **2018**, *9*, 293.
- (83) Gao, W.; Gao, X.; Abtey, T. A.; Sun, Y.-Y.; Zhang, S.; Zhang, P. *Phys. Rev. B* **2016**, *93*, 085202.

- (84) Umadevi, D.; Watson, G. W. *ACS Omega* **2019**, *4*, 5661–5669, PMID: 31459720.
- (85) Vona, C.; Nabok, D.; Draxl, C. *Advanced Theory and Simulations* **2022**, *5*, 2100496.
- (86) Hybertsen, M. S.; Louie, S. G. *Physical Review B* **1986**, *34*, 5390–5413.
- (87) Deslippe, J.; Samsonidze, G.; Strubbe, D. A.; Jain, M.; Cohen, M. L.; Louie, S. G. *Computer Physics Communications* **2012**, *183*, 1269–1289.
- (88) Guandalini, A.; D’Amico, P.; Ferretti, A.; Varsano, D. *npj Computational Materials* **2023**, *9*, 44.
- (89) Marini, A.; Hogan, C.; Grüning, M.; Varsano, D. *Computer Physics Communications* **2009**, *180*, 1392–1403.
- (90) Bruneval, F.; Gonze, X. *Phys. Rev. B* **2008**, *78*, 085125.
- (91) Govoni, M.; Galli, G. *Journal of Chemical Theory and Computation* **2015**, *11*, 2680–2696, PMID: 26575564.
- (92) Kresse, G.; Furthmüller, J. *Computational Materials Science* **1996**, *6*, 15–50.
- (93) Debela, T. T.; Yang, M. C.; Hendon, C. H. *Journal of the American Chemical Society* **2023**, *145*, 11387–11391.
- (94) Li, C.; Zhang, L.; Chen, J.; Li, X.; Sun, J.; Zhu, J.; Wang, X.; Fu, Y. *Nanoscale* **2021**, *13*, 485–509.
- (95) Hendon, C. H.; Rieth, A. J.; Korzyński, M. D.; Dincă, M. *ACS Central Science* **2017**, *3*, 554–563.
- (96) Syzgantseva, M. A.; Stepanov, N. F.; Syzgantseva, O. A. *ACS Applied Materials and Interfaces* **2020**, *12*, 17611–17619.
- (97) Hua, C.; Doheny, P. W.; Ding, B.; Chan, B.; Yu, M.; Kepert, C. J.; D’Alessandro, D. M. *Journal of the American Chemical Society* **2018**, *140*, 6622–6630.
- (98) Xie, L. S.; Skorupskii, G.; Dincă, M. *Chemical Reviews* **2020**, *120*, 8536–8580.
- (99) Wang, H.-F.; Chen, L.; Pang, H.; Kaskel, S.; Xu, Q. *Chemical Society Reviews* **2020**, *49*, 1414–1448.
- (100) Jin, S. *ACS Energy Letters* **2019**, *4*, 1443–1445.
- (101) Liu, J.; Zhu, D.; Guo, C.; Vasileff, A.; Qiao, S. *Advanced Energy Materials* **2017**, *7*.

- (102) Zhang, H.; Liu, X.; Wu, Y.; Guan, C.; Cheetham, A. K.; Wang, J. *Chemical Communications* **2018**, *54*, 5268–5288.
- (103) Tahir, M. A.; Arshad, N.; Akram, M. *Journal of Energy Storage* **2022**, *47*, 103530.
- (104) Sheberla, D.; Bachman, J. C.; Elias, J. S.; Sun, C.-J.; Shao-Horn, Y.; Dincă, M. *Nature Materials* **2016**, *16*, 220–224.
- (105) Maiti, S.; Pramanik, A.; Dhawa, T.; Sreemany, M.; Mahanty, S. *Materials Science and Engineering: B* **2018**, *229*, 27–36.
- (106) Park, J. G.; Aubrey, M. L.; Oktawiec, J.; Chakarawet, K.; Darago, L. E.; Grandjean, F.; Long, G. J.; Long, J. R. *Journal of the American Chemical Society* **2018**, *140*, 8526–8534.
- (107) Diamond, B. G.; Payne, L. I.; Hendon, C. H. *Communications Chemistry* **2023**, *6*.
- (108) Hmadeh, M. et al. *Chemistry of Materials* **2012**, *24*, 3511–3513.
- (109) Park, G.; Demuth, M. C.; Hendon, C. H.; Park, S. S. *Journal of the American Chemical Society* **2024**,
- (110) Apostol, P.; Gali, S. M.; Su, A.; Tie, D.; Zhang, Y.; Pal, S.; Lin, X.; Bakuru, V. R.; Rambabu, D.; Beljonne, D.; Dincă, M.; Vlad, A. *Journal of the American Chemical Society* **2023**,
- (111) Lu, Y. et al. *Nature Communications* **2022**, *13*.
- (112) Day, R. W.; Bediako, D. K.; Rezaee, M.; Parent, L. R.; Skorupskii, G.; Arguilla, M. Q.; Hendon, C. H.; Stassen, I.; Gianneschi, N. C.; Kim, P.; Dincă, M. *ACS Central Science* **2019**, *5*, 1959–1964.
- (113) Nam, K. W.; Park, S. S.; dos Reis, R.; Dravid, V. P.; Kim, H.; Mirkin, C. A.; Stoddart, J. F. *Nature Communications* **2019**, *10*.
- (114) Clough, A. J.; Skelton, J. M.; Downes, C. A.; de la Rosa, A. A.; Yoo, J. W.; Walsh, A.; Melot, B. C.; Marinescu, S. C. *Journal of the American Chemical Society* **2017**, *139*, 10863–10867.
- (115) Sheberla, D.; Sun, L.; Blood-Forsythe, M. A.; Er, S.; Wade, C. R.; Brozek, C. K.; Aspuru-Guzik, A.; Dincă, M. *Journal of the American Chemical Society* **2014**, *136*, 8859–8862.
- (116) Skorupskii, G.; Trump, B. A.; Kasel, T. W.; Brown, C. M.; Hendon, C. H.; Dincă, M. *Nature Chemistry* **2019**, *12*, 131–136.

- (117) Skorupskii, G.; Dincă, M. *Journal of the American Chemical Society* **2020**, *142*, 6920–6924.
- (118) Nath, A.; Kumar, V.; Shukla, A.; Ghosh, H. N.; Mandal, S. *Angewandte Chemie International Edition* **2023**, *62*.
- (119) Brédas, J.-L.; Beljonne, D.; Coropceanu, V.; Cornil, J. *Chemical Reviews* **2004**, *104*, 4971–5004.
- (120) Hsu, C.-P. *Accounts of Chemical Research* **2009**, *42*, 509–518.
- (121) Demuth, M. C.; Hendon, C. H. *ACS Materials Letters* **2023**, *5*, 1476–1480.
- (122) Sun, L.; Park, S. S.; Sheberla, D.; Dincă, M. *Journal of the American Chemical Society* **2016**, *138*, 14772–14782.
- (123) Schimka, L.; Harl, J.; Kresse, G. *The Journal of Chemical Physics* **2011**, *134*.
- (124) Frisch, M. et al. *Gaussian Inc. Wallingford CT* **2009**,