Quantum Chemistry in Solid-State Simulations: Gaussian Basis Sets Development and ab-initio Green's Function Based Realistic System Applications

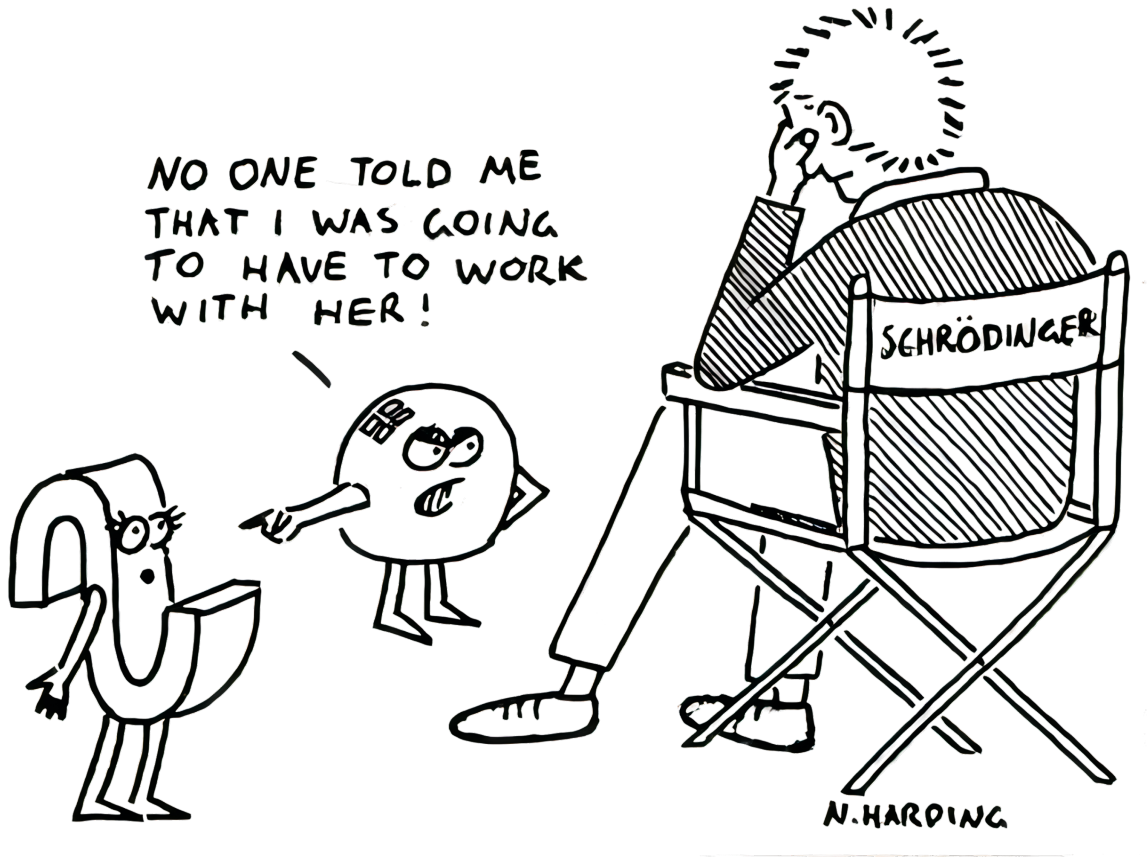
by

Yanbing Zhou

A dissertation submitted in partial fulfillment
of the requirements for the degree of
Doctor of Philosophy
(Chemistry and Scientific Computing)
in The University of Michigan
2022

Doctoral Committee:

Associate Professor Dominika Zgid, Chair
Professor Julie Biteen
Professor Vikram Gavini
Professor Eitan Geva


NO ONE TOLD ME THAT I WAS GOING TO HAVE TO WORK WITH HER!
SCHRÖDINGER
N.HARDING

Yanbing Zhou
yanbing@umich.edu
ORCID iD: 0000-0001-7673-771X

DEDICATION

To my family, for their never ending support.

ACKNOWLEDGEMENTS

In this part of my dissertation, I would like to express my thanks and appreciation to many people, without whose support, I would not have gone through the journey from where I started to where I am now.

First and foremost, I would like to extend my deepest gratitude to my research advisor Dr. Dominika Zgid, who has opened the door of quantum chemistry for me, and guided me throughout my degree. As a knowledgeable scientist, her guidance and support have shaped me into a scientific researcher who entered the group with no relevant background in the field of study. Her ample knowledge of physical chemistry and quantum simulations, scientific intuition, and problem-solving skills are essential in the progress of this dissertation. And as a caring mentor, her patience and encouragement are invaluable to me in the frustrating moments. She gave me all the freedom I needed while guiding me on the best path forward and broadened my vision in the field by providing various opportunities for me to attend conferences and communicate with preeminent domestic and international scientists. Words cannot do justice to how grateful I am for what she offered over the past years.

I am incredibly grateful to Dr. Emanuel Gull for his co-supervising of the projects listed in this dissertation. Without the enormous valuable insight and inspiring advice he continuously provided, the process of completing this work would not be as smooth.

I would like to thank Dr. Eitan Geva for his mentorship. I learned greatly from Eitan during my rotation in his group and from the quantum dynamics coursework I took with him. These experiences provided me with a solid and profound under-

standing of the relevant field.

I would like to thank my committee members Dr. Julie Biteen, Dr. Vikram Gavini and Dr. Eitan Geva, for their advice, critical feedback, and time during meetings along the path to my degree.

I would like to thank the members of Zgid's group, past and present: Dr. Alexander Rusakov, from whom I have learned a tremendous amount of solid-state knowledge; Dr. Tran Nguyen Lan, for his helpful advice given at the time I joined the group; Dr. Alicia Welden for the guidance and direction that paved my way to find an excellent position after graduation; Dr. Alexei Kananenka, Dr. Avijit Shee and Yuting Chen for always being supportive colleagues; and at last, Dr. Blair Winograd and Diksha Dhawan, for all the friendship, accompany, encouragement and time we spent working together. I appreciate the moments I spent discussing sciences with them.

I was fortunate to have incredible collaborators from the Gull group, from whom I have learned and grown as a researcher. I would like to thank Dr. Sergei Iskakov for his mentorship in the condensed-matter field and friendship. And I would like to thank Chia-nan Yeh and Runxue Yu for all the discussions we had on the SEET application project.

Many thanks to the friends I have in the Chemistry and Physics department who made grad school a more delightful place to be in, and with whom I enjoyed discussing sciences and having random chats, including Hanjie Jiang, Alan Rask, Yifan Lai and Xingyang Dong. Special thanks to my dear friend Rong. Your accompany and encouragement throughout my years in Michigan are meaningful.

Last but not least, I want to thank my family for their support all this time. This journey wouldn't even begin without them.

TABLE OF CONTENTS

LIST OF FIGURES

Figure

LIST OF TABLES

Table

LIST OF ABBREVIATIONS

AO atomic orbital

cGTOs contracted Gaussian-type orbitals

CISD configuration interaction singles and doubles

DFT density functional theory

ED Exact Diagonalization

GF2 second-order Green's function theory

GGA generalized gradient approximations

GTOs primitive Gaussian-type orbitals

HF Hartree-Fock

LAPW linear augmented plane waves

LDA local density approximation

MO molecular orbital

MP2 Møller-Plesset perturbation theory

SAO symmetrized atomic orbitals

scGW self-consistent GW approximation

ABSTRACT

Understanding emergent many-body phenomena in correlated materials remains one of the grandest challenges in solid-state quantum chemistry calculations. The individuality of electrons diminishes as correlation increases, which leads to the emergence of new features in the system. Seeking to address such problems accurately requires the aid of non-perturbative many-body techniques as well as the corresponding high-quality basis sets. In periodic systems, a given element may be present in different spatial arrangements displaying vastly different physical and chemical properties, and an elemental basis set independent of the physical properties of materials may lead to significant simulation inaccuracies. In fact, with the rapid progress of quantum chemistry methods in condensed-phase simulations, the need for a library of reliable Gaussian basis sets explicitly designed for periodic calculations has become urgent. In the meantime, the development of parameter-free, systematic and reliable quantum chemistry methods for simultaneously treating weak and strong correlations in periodic systems has never stopped. Their ability to treat realistic solid materials needs to be tested, and the obtained results can serve as benchmarks for empirical parameter-guided solid-state calculations.

In this thesis, we present numerical studies of various solid-state systems, emphasizing the design and optimization of Gaussian basis sets and assessing the band structure properties of the periodic solid systems. The work begins by introducing the theoretical framework of the electronic structure theory and presenting the fundamentals of a fully self-consistent GW approximation and the self-energy embedding theory (SEET) in Chapters II and III. It continues with a description of the Gaussian

basis set optimization scheme we devised for solid-state quantum chemistry calculations. The scheme we present is designed to avoid a lack of material specificity within a given basis set by simultaneously minimizing the total energy of the system and optimizing the band energies when compared to the reference plane wave calculation while accounting for the overlap matrix condition number. We compare the quality of the Gaussian basis sets generated via our method against the existing basis sets. And the optimization scheme is tested to yield improved results.

We finally present a quantitative study of electronic properties of realistic materials with the treatment of Green's function based weakly correlated self-consistent GW approximation and the strongly correlated self-energy embedding theory. In the transition metal oxide BiVO3, we found the inaccurate illustration of orbitals including t_{2g} in transition metal Vanadium by DFT. Our calculation shows a systematic trend of DFT failure in illustrating both insulating and metallic solutions, demonstrating the dangers of using DFT when the experimental data is scarce.

CHAPTER I

Introduction

The exploration and synthesis of potentially useful materials for modern day technology can be costly and time-consuming using only experimental techniques. However, this process can be actively supported and facilitated by computational modeling. Making testable predictions on properties of realistic materials requires a theoretical model that can accurately capture the relevant physics and simultaneously is computationally feasible to obtain a solution. A reasonable choice of electronic structure methods and a high-quality Gaussian basis set is always at the core of an effective and accurate quantum chemistry periodic simulation.

The many fascinating phenomena in solids, such as Magnetism, Mott transitions, Kondo and heavy-fermion behavior, and superconductivity, arises from the motions and many-body interaction of electrons in the system. However, it is also this electron-electron interaction that makes modeling and seeking solutions for such systems challenging. With the extensive system size and a large number of electrons in the periodic solids, solving the many-electron Hamiltonian conventionally with wave function based methods is not feasible, as the number of Slater determinants required for an accurate solution is significant. Alternatively, with its simple formalism and its ability to treat systems with hundreds of atoms, Density-functional theory (DFT) is considered the Standard Model for solid-state calculations. The success of DFT

in explaining the physical and chemical properties of solids is considered remarkable, however, the approximation to DFT, the local density approximation (LDA), and its simple extensions fails for the systems in which the individuality of electrons diminishes. In other words, standard DFT approximations face difficulty capturing electron correlations in the solid systems, leading to quantitative errors in band gaps prediction for insulators and semiconductors, especially for those with localized d shell electrons. Furthermore, qualitative failures exist occasionally. For instance, Ge is incorrectly predicted to be a semi-metal by LDA, while experimentally it is determined to be a semi-conductor. The introduction of dynamical mean-field theory (DMFT) combined with DFT is a remedy to this issue. However, this is realized at the price of inducing adjustable parameters into the formalism of the method. Alternatively, diagrammatic perturbation methods free from empirical parameters can be considered. One of the most fruitful and well-known methods among them is the GW approximation. Its success is phenomena in band gaps and spectral properties prediction of solids. However, its perturbative nature leads to its failure in the capturing of strong electronic correlation.

The search for a rigorous parameter-free ab-initio theory for simultaneously treating weakly and strongly correlated systems has always been an ongoing task. The goal is to find approximate methods that could accurately treat the electronic correlation while reasonably capturing the physics of the remaining system from a quantum chemical perspective. In this work, we introduce and assess a diagrammatic ab initio embedding theory: the self-energy embedding theory (SEET), which in combination with the weakly correlated GW approximation, can give a reliable description of solid systems with localized electrons involved. Transition metal oxides play vital roles in many energy-related problems. In practice, various schemes were applied to such systems with inevitable success or failures, yet a universal tool for treating such systems has not been found. For those transition metal oxides not being well characterized

experimentally, a parameter-free ab-initio becomes rather crucial. And considering the correlated nature of transition metal d orbitals, such systems, therefore, can serve as an ideal testbed for SEET.

In the meantime, with the rapid progress of quantum chemistry methods in condensed-phase simulations, it becomes demanding for quantum chemists to design and optimize Gaussian basis sets for periodic calculations. In fact, the plane wave basis was far more commonly used than any other type of basis sets for periodic systems with delocalized electrons for many reasons, including its natural periodicity, systematic convergence property, and the reduction of Hamiltonian second-quantized terms in the plane wave representation. However, plane wave calculations are only affordable due to the low computational scaling of DFT, as thousands of plane waves are necessary to reach a basis set accuracy comparable to that of a Gaussian basis set with many fewer basis functions for most materials. For post-DFT ab-initio calculations with a higher computational scaling, Gaussian basis sets become more appealing.

Most of the existing Gaussian basis sets in quantum chemistry are made for atoms and molecular simulations. Their application to close-packed solids is problematic, as the basis sets are quickly getting linearly dependent. Also, due to the nature of solids, a given element may display vastly different physical and chemical properties in different spatial arrangements.

The rest of the dissertation is as follows. Chapter II and III introduce the numerical algorithms of the electronic structure theory relevant to this work, with the former one focusing on DFT while the latter one on Green's function based correlated methods GW and SEET. Chapter IV present and examine a material-specific Gaussian basis optimization scheme for solids. Chapter V report results from a fully self-consistent ab initio SEET calculation applied to a transition metal oxide system. And chapter VI concludes this thesis.

CHAPTER II

Background: Mean-field Methods

2.1 Mean-field Approximation

Understanding the properties of materials requires confronting the many-body problem. The problem with calculating systems composed of interacting particles exactly comes when one attempts to enumerate all the possible microstates and calculate their energy. While at first, this appears to be an impossible task, substantial progress has been made by combining physical insights with modern numerical approaches. The key strategy is to map the complex realistic many-body Hamiltonian to a simpler auxiliary model that can be solved numerically. In computational chemistry, one of the commonly adopted and fundamental ways of picturing the electronic structure of materials is to use the mean-filed model. In general, the mean-field theory allows the replacement of an N-body system by a 1-body system that sits in a suitably chosen external field. The external field is set equal to the average field due to the remaining particles. Hence, this approximation technique allows one to map a multi-body problem onto a one-body problem.

2.1.1 Electron Correlation

To chemical systems specifically, the mean-filed theory is treating every single electron in the system of studying as they are moving in a mean-field formed by other

electrons in the system. As a result, (for electronic structure methods which deploy the mean-field approximation like Hartree-Fock), the repulsion energy between two electrons is calculated between an electron and the average electron density formed by the other electrons. This type of approximation gives rise to the problem of electron-electron repulsion overestimation due to the fact that it doesn't take into account the instant motions of the other electrons when the selected electron approaches. This instant motion diminished the repulsion energy, and this concept of electrons avoiding each other is defined as "dynamical correlation". The electronic energy quantity that arises from this dynamic correlation is called the dynamic correlation energy.

While this energy appears to be small in quantity relative to the total electronic energy of the system, it is crucial to include a certain level of description of it in the calculation, as it may be crucial to the chemical properties of the system of interest. In fact, the inclusion of correlation energy is especially important for solid-state systems with closely packed electrons, as a pure mean-filed treatment of such systems may give quantitative, or even qualitative wrong predictions. We have seen these examples in the mean-field calculations like Hartree-Fock approximations (HF) and Density Functional Theory (DFT), while both of them are well-known schemes designed to model the properties of many-electron systems, the latter one is, in particular, most commonly employed by the solid-state community, especially for the ground-state total energies and charge densities. The next section of this chapter will explore the fundamentals of DFT and its applications to crystalline solids.

2.2 Many-particle Problems in Solids

The quantum states of an N-particle isolated system can be described by a wave function: $\psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N)$, with $\mathbf{r}_i$ representing the particle coordinate in three-

dimensional space. By solving the electronic Schrödinger equation:

$$\hat{H}_e\psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) = E\psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) \tag{2.1}$$

, the ground state of a collection of atoms in a crystalline solid can be found. The electronic Hamiltonian of this N-particle system, with neglecting the the kinetic energy of the nuclei and the repulsion between the nuclei of the system, can be written in three terms:

$$\hat{H}_e = -\frac{1}{2}\sum_{i=1}^{N}\nabla_i^2 - \sum_{i=1}^{N}\sum_{A=1}^{M}\frac{Z_A}{|\mathbf{r}_i - \mathbf{R}_A|} + \sum_{i<j}\frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} \tag{2.2}$$

, with $\mathbf{R}_A$ and Z_A be the nuclear positions and charges in the system.

This is based on the Born-Oppenheimer approximation, the central concept of quantum chemistry, which treats the nuclei in the system as fixed external potential faced by moving electrons, hence decoupling the dynamics of the two types of particles.

The first term of (2.3) is known to be the operator for the kinetic energy of the N-electrons; the second term represent the ionic potential between nuclei and electrons; and the third term represent the pair-wise electronic repulsion. Hence, the electronic Hamiltonian can be written as following:

$$\hat{H}_e = \hat{T}_{\mathbf{r}} + \hat{V}_{ext} + \hat{V}_U. \tag{2.3}$$

For most of the crystalline systems, the computational complexity required to solve (2.1) is beyond the reach of current technologies. To illustrate the dimension of the Schrodinger equation, consider of solving the ground state energy of a simple solid-state systems, for instance, silicon with 14 electrons per atom, and a total of 100 atoms in the system. With three spatial coordinate per electronic position, the problem

becomes 4200 dimensional. And for some slightly more complicated materials, this number can soon reach 1e4-1e5. It is quite obvious that the exact many-body problem becomes dramatically expensive as soon as the number of electrons in the system increases. Computational chemists and physicists hence started to consider replacing the highly complicated many-body wave functions, with a simple, three-dimensional scalar function, the electronic density. This leads to the concept of density functional theory (DFT) , one of an exact mean-field formalism for calculating the ground-state properties of a many-electron system. The central idea of DFT lies on the fact that there is no need to know the details of the N-particle wave functions in order to calculate the ground-state properties. In fact, the knowledge of the electronic density, $n(r)$, is enough to obtain the ground state energy of the system and all quantities that are related.

2.3 DFT and Kohn-Sham Equation

DFT is based on two fundamental theorems, which states that there is a one-to-one correspondence between the ground-state density $n_0(r)$ and the external potential as well as a variational principle for energy functional $E[n_0] \leq E[n]$. This is known as the Hohenberg-Kohn theorem. The first statement in the Hohenberg-Kohn theorem implies that the many-particle Hamiltonian is a functional of the ground-state density. The latter statement allows one to obtain the ground state of a many-electron system by the variation of its density, hence reducing the 3N-dimensional problems to a 3-dimensional problem. The theorem stated above can be generalized to include spin. The electronic Hamiltonian and the energy functional of the system can be written as:

$$\hat{H}_e = \hat{T}_{\mathbf{r}} + \int v_{ext} n_(\mathbf{r}) d\mathbf{r} + \hat{V}_U \tag{2.4}$$

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

With $v_{ext}(\mathbf{r})$, the second term in (2.5) is system dependent, which the rest of the terms in the equation are universal. The fist term of (2.5) consists of 2 contributors, a classical Coulomb interaction term and a non-classical self-interaction correction term:

$$\int \frac{n(\mathbf{r})n(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} d\mathbf{r}d\mathbf{r}' + E_{xc}^{true} \tag{2.6}$$

and latter one is the so-called exchange-correlation term that requires approximation in practice.

In 1965, Kohn and Sham introduced their new scheme of obtaining the ground state electron density $n_0(\mathbf{r})$ that includes a transformation of the original N-particle system into an non-interacting auxiliary particles in effective potential:

$$\psi(\mathbf{r}_1, \mathbf{r}_2, \ldots, \mathbf{r}_N) = \psi_1(\mathbf{r}_1)\psi_2(\mathbf{r}_2)\ldots\psi_N(\mathbf{r}_N) \tag{2.7}$$

The Kohn-Sham equations are a set of N Schrodinger-like one-particle equations with a modified effective potential:

$$(-\frac{1}{2}\nabla^2 + v_{ext}(\mathbf{r}) + v_H(\mathbf{r}) + v_{xc}(\mathbf{r}))\psi_i(\mathbf{r}) = \epsilon_i(\mathbf{r})\psi_i(\mathbf{r}) \tag{2.8}$$

. The interaction of the system are hence accounted in the potential terms, with a contribution from the quantum mechanical exchange-correlation term. Notice that the explicit expression for $v_{xc}(\mathbf{r}))$ is inaccessible, and it is relatively small in quantity with respect to the kinetic and coulomb term in (2.8). Although small in quantity, however, the XC term plays a central role as it directly affects the accuracy and efficiency of the theory. Solving N one-particle (2.8) is obviously easier than the original 3N-dimensional problem. And the total energy of this N one-particle system

with effective potential can be expressed as:

$$E[n(\mathbf{r})] = T_0[n(\mathbf{r})] + \int v_{ext} n_(\mathbf{r}) d\mathbf{r} + E_H[n(\mathbf{r})] + E_{xc}[n(\mathbf{r})] \tag{2.9}$$

. In practice, the construction of an increasingly exact approximations to exchange-correlation functionals $E_{xc}[n(\mathbf{r})]$ is still an on-going challenge within DFT. The next section will explore some of the existing successes and some central considerations.

2.4 Exchange-Correlation Functionals

As stated in the Sec. 2.3, DFT does not explicitly describe how the gound-state total energy depends on the ground-state electronic density. And this inexactness is due to the unknown of the exact functional in the Hohenberg-Kohn theorem. In practice, approximations are needed to express the kinetic term and the many-body electron-electron interaction term as a functional of $n(\mathbf{r})$. One thus have to leave the exact DFT formalism, and reaches the complicated side of approximations with their specific range of validity.

It is worth noting that, however, DFT can deliver reasonable results with certain approximations for certain functionals in certain scenatios, even includes the LDA. And this accuracy in combination with the simplicity of DFT formalism, which allows systems with several hundred atoms to be calculated, leads to the phenomenal success of DFT in terms of the number of users and systems studied.

Proposed by Hohenberg and Kohn in 1964, the LDA relies on the approximation assuming that the exchange-correlation energy for an electron located in spacial location $\mathbf{r}$ only depends on the local value of the electronic density $n(\mathbf{r})$. Such an approximation is strictly valid only in the limit of a homogeneous electron gas. Hence, LDA tends to be more reliable on 3-dimensional periodic systems than finite systems with less homogeneous electron density distribution. In the paper (*Hohenberg and*

Kohn, 1964b), Hohenberg and Kohn concluded that chance of LDA being successful on realistic systems with strong variations of charge density is small. However, in practice, LDA has demonstrated remarkable success: the accuracy is within 1% with respect to experimental data for certain types of observables like lattice parameters on certain non-metallic systems like diamonds and silicons. And when it comes to band gap prediction, LDA result are mostly giving qualitatively correct, while quantitatively they are less reliable, especially for insulators and semiconductors. Band gaps in those systems are often being underestimated approximately 1.0 eV or more. For instance, LDA incorrectly predicts Ge to be a semi-metal (with a negative band gap) instead of a semiconductor (*Csonka et al.*, 2009; *Aulbur et al.*, 2000).

In practice, it remains an central issue in DFT field to improve from the LDA approximation. There exist two main trend of functional development. The first requires the functionals to satisfy exact mathematical relations like XC hole sum rule, while the second path consists of fitting functional parameters against experimental data.

For the former path, one practice is to account for the charge inhomogeneities in true electron density, and this includes considering not only the local density but also its gradient. The information of inhomogeneity allows corrections based on the changes in density away from the coordinate. Since mid-1980s, the so-called generalized gradient approximations (GGA) started to become popular and substantially increased the accuracy of DFT. The GGA family consists of some well-known functionals, like the Becke's exchange functional B88 (*Becke*, 1988), the Perdew–Burke–Ernzerhof's exchange correlation functional PBE (*Perdew et al.*, 1997), the Perdew–Wang's PW91 (*Perdew et al.*, 1992) and the Lee–Yang–Parr's LYP (*Lee et al.*, 1988) correlation functionals. And as result, PBE, for instance, leads to better predictions than LDA on metallic systems, but shows no significant improvement for non-metallic systems (*Perdew et al.*, 2009; *Csonka et al.*, 2009).

With the well-known existing limitations of LDA and GGA functionals as discussed above, a new approach has been proposed, and in which HF and DFT are being combined and mixed. These so-called hybrid functionals are in principle a form of DFT generalization, and are nowadays most effective and popular funcitonals.

In hybrid functionals, $E_{xc}[n(\mathbf{r})]$ consists of a portion of one-body-orbital dependent exact exchange. This comes from the fact that the total exchange-correlation energy are predominantly contributed by the exchange energy. In practice, it is common to dive E_{xc} into seperate functionals:

$$E_{xc} = E_x + E_c \tag{2.10}$$

, namely the exchange and correlation terms, and each of them are modeled separately. It is therefore reasonable for one to consider replacing the local and semilocal E_x in LDA and GGA, with an essentially non-local exact exchange energy (HF), which includes only the Fermi electron correlation with identical spins.

Due to this additional cost from HF, this type of functionals are more expensive than pure density funcitonals. The advantages of hybrid functionals, however, stems from the fact that HF tends to overestimate gaps due to the lack of correlations as discussed in the previous section. This is completely on the contrary to Kohn-Sham DFT result. The mixing of certain amount of exact exchange with DFT XC potentials hence leads to better gaps.

This has been exemplified with B3LYP functional, which includes 20% ($a_0 = 0.2$) of exact exchange (*Becke*, 1993), which provides better description accuracy than some of the non-hybrid functionals. And this successfulness stimulated the development of new hybrid functionals. It is worth noting that numerical parameters taken in E_{xc}^{hybr} are obtained by fitting computational results against experimental data. In practice, the criteria for selecting the proper amount of exact exchange for a

given system is therefore challenging if one wishes to preserve an ab initio approach. Also, due to this experimental data dependent functional parameters fitting scheme, developing or choosing appropriate DFT functional may becomes problematic for computational researchers when there is a lack of enough empirical guidance.

2.5 Limitations

It is very important to reiterate the fact that DFT calculations in practice are not exact solutions of the full Schrödinger equation. There is an intrinsic uncertainty that exists between the energies calculated with DFT and the true ground-state energies of the Schrödinger equation from time to time. Therefore, one should not expect DFT to be physically accurate in all scenario.

The first situation where DFT calculations have limited accuracy is in the calculation of electronic excited states, and this is majorly because DFT is restricted to ground-state properties in its original formulation.

Band structure calculation, with the need of electronic energy levels information, is therefore beyond the reach of standard DFT. In practice, it is common to exploit Kohn-Sham eigenvalues, the solutions of the N one-particle equations with effective potential (2.8), for band structure plot generation. The solution of the Kohn-Sham eigenvalue equation yields a whole spectrum of single-particle states. One may be tempted to identify the corresponding eigenvalues with excitation energies, however, the rationale for such correspondence has not been firmly established. This is partially due to the fact that the Kohn-Sham eigenvalues are better to be interpreted as mathematical tools (Lagrange multipliers) without an actual physical meaning. Additionally, the sum of Kohn-Sham eigenvalues over occupied energy states is different from the ground-state total energy of an N-electron system. Consequently, the band gap prediction of Kohn-Sham DFT is known to suffer from certain level of failures. One example would be the LDA underestimation of band gaps in semiconducting

and insulating materials, such error can go up to 2 factors with comparing to the experimental data. Another well-known failure of LDA is in the case of germanium. While experimentally measured to be semiconductor, Ge was computationally gauged as semi-metal. This failure of band gap prediction can relate to the lack of sensitivity of local charge density with respect to the subtraction or addition of a electron (electron delocalized). This gap estimation tendency is supported by results presented Chapter V as well, a detailed discussion can be seen in the later chapter.

Other limitations of DFT includes the many-electron self-interaction error, which can be eliminated by using exact exchange functional. These intrinsic deficiencies restrain the accuracy of calculations on realistic solids. Qualitative, and sometimes quantitative failures exist and are not limited to the cases discussed above. Among a number of documented situations, one of the most prominent failure of DFT is on its treatment of strongly correlated materials, like transition metal oxides (*Imada et al.*, 1998).

CHAPTER III

Green's Function Based Correlated Methods

While ground state DFT could tremendously help in gaining physical and intuitive insight for weakly correlated systems, more explicit methods are occasionally needed in order to correct them, especially when the electronic correlation of the system becomes strong. For certain observable like spectra, explicit expressions in terms of density or the density matrix are unknown. And to merge time and temperature information into the solution, it is natural for Green's function formalism, an elegant alternative to DFT, to be considered.

The primary development of Green's function theory can be traced back to 1950's and early 60's. In 1954, the Lehmann representation was devised by *Lehmann* (1954). A summarize of the perturbative approach to Green functions can be found in *Abrikosov et al.* (1975). In the 60's, a functional approach was developed by Joaquin Luttinger and John Ward (*Luttinger and Ward*, 1960) , while Matsubara presented a temperature-dependent formalism of Green's function in 1955 (*Matsubara*, 1955).

We present the Green's function based correlated electronic structures methods, GW approximation, and on top of it, the quantum embedding method SEET, in the following chapter.

3.1 GW approximation

3.1.1 One-particle Green's Function in Second-Quantized Form

The one-body Green's function is an efficient mathematical tool that describes electron addition or subtraction to or from a system, and is defined as:

$$\begin{aligned} G(\mathbf{r},t;\mathbf{r}',t') &= -i\,\langle\Psi|T[\hat{\psi}(\mathbf{r},t)\hat{\psi}^{\dagger}(\mathbf{r}',t')]|\Psi\rangle \\ &= -i\,\langle\Psi|\hat{\psi}(\mathbf{r},t)\hat{\psi}^{\dagger}(\mathbf{r}',t')|\Psi\rangle\,\theta(t-t') + i\,\langle\Psi|\hat{\psi}^{\dagger}(\mathbf{r}',t')\hat{\psi}(\mathbf{r},t)|\Psi\rangle\,\theta(t'-t) \end{aligned} \tag{3.1}$$

where T is the Wick time-ordering operator, which rearranges operators in ascending time arguments from right to left:

$$T[\hat{\psi}(\mathbf{r},t)\hat{\psi}^{\dagger}(\mathbf{r}',t')] = \begin{cases} \hat{\psi}(\mathbf{r},t)\hat{\psi}^{\dagger}(\mathbf{r}',t') & \text{for } t > t' \\ \hat{\psi}^{\dagger}(\mathbf{r}',t')\hat{\psi}(\mathbf{r},t) & \text{for } t' > t \end{cases} \tag{3.2}$$

and $\theta(t)$ is the Heaviside step function:

$$\theta(t) = \begin{cases} 1 & \text{for } t > 0 \\ 0 & \text{for } t < 0 \end{cases} \tag{3.3}$$

, $\psi^{\dagger}$ and ψ are the creation and annihilation field operators in the Heisenberg representation, respectively. Notice that the spin dependence are not taken account explicitly here, however, it can be taken as part of the spatial coordinate $\mathbf{r}$ if necessary.

Following Eq. 3.1, the electron density of the system can be expressed in terms of G via:

$$n(\mathbf{r},t) = \langle\Psi|T[\hat{\psi}(\mathbf{r},t)\hat{\psi}^{\dagger}(\mathbf{r}',t')]|\Psi\rangle = -i\hbar G(\mathbf{r},t;\mathbf{r},t+\eta) \tag{3.4}$$

, with η be an infinitesimal positive number. It helps to enforce the proper time

ordering of the field operator. The Fourier transform of $G(\mathbf{r}, t; \mathbf{r}', t')$ $(t - t' \to \omega)$ to the frequency space using the Fourier transform of the Heaviside side function takes the Lehmann representation below:

$$G(\mathbf{r}, \mathbf{r}'; \omega) = \sum_i \frac{\psi_i^{N+1}(\mathbf{r})\psi_i^{N+1^*}(\mathbf{r}')}{\hbar\omega - \epsilon_i^{N+1} + i\eta} + \frac{\psi_i^{N-1}(\mathbf{r})\psi_i^{N-1^*}(\mathbf{r}')}{\hbar\omega - \epsilon_i^{N-1} - i\eta} \tag{3.5}$$

It is observed that pole exist in $G(\mathbf{r}, \mathbf{r}'; \omega)$ at the many-particle exitation energy $\epsilon i^{N\pm1}$. And ϵi^{N+1} and ϵi^{N-1} correspond to the excitation energy measured in the inverse and direct photoemission spectroscopy.

Since Eq. 3.5 requires the knowledge of many-body wavefunctions that are hard to calculate and store in systems with more than few electrons, it appears to be less useful in practice. Following from that, an alternative class of approach has been proposed. It can be shown that $G(\mathbf{r}, \mathbf{r}'; \omega)$ obeys the Dyson equation:

$$G(\mathbf{r}, \mathbf{r}'; \omega) = G_0(\mathbf{r}, \mathbf{r}'; \omega) + \int\int G_0(\mathbf{r}, \mathbf{r}''; \omega)\Sigma(\mathbf{r}'', \mathbf{r}'''; \omega)G(\mathbf{r}''', \mathbf{r}'; \omega)d^3\mathbf{r}''d^3\mathbf{r}''' \tag{3.6}$$

where G_0 is the Green's function of a non-interacting system, and the non-local Σ is frequency dependent, and is the so-called self-energy that describes the scattering event.

As an type of single-particle propagator, the one-body Green's function can be interpreted as the probability amplitude that detects either: an electron at $(\mathbf{r}, t)$ when an electron is injected to the system at $(\mathbf{r}', t')$, or an hole at $(\mathbf{r}, t)$ when an electron is deleted from the system at $(\mathbf{r}', t')$. Since the propagation process can includes different order, the Green's function is defined as the summation of the possibility amplitude for each of the possible particle interaction process. This leads to a typical equation of scattering theory:

$$G = G_0 + G_0\Sigma G_0 + G_0\Sigma G_0\Sigma G_0 + G_0\Sigma G_0\Sigma G_0\Sigma G_0 + \cdots \tag{3.7}$$

where the terms in Eq. 3.7 describes zero, single, double, triple, etc., scattering process.

This geometric series expression is derived from Eq. 3.6, and can be rewritten into a symbolic form:

$$G = G_0 + G_0 \Sigma G \tag{3.8}$$

By inserting Eq. 3.5 into Eq. 3.6 to obtain the quasi-particle equation:

$$(-\frac{1}{2}\nabla^2 + v_{ext}(\mathbf{r}) + v_H(\mathbf{r}))\psi_i(\mathbf{r}) + \underbrace{\int \Sigma(\mathbf{r}, \mathbf{r}'; \epsilon_i)\psi_i(\mathbf{r})d^3\mathbf{r}'}_{\text{Approximations}} = \epsilon_i(\mathbf{r})\psi_i(\mathbf{r}) \tag{3.9}$$

where the Coulomb interaction is constituted by a classical electrostatic Hartree potential:

$$v_H = \int \frac{n(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|} d^3\mathbf{r}' \tag{3.10}$$

Recall from Chapter II that the KS equation takes the following form:

$$(-\frac{1}{2}\nabla^2 + v_{ext}(\mathbf{r}) + v_H(\mathbf{r}))\psi_i(\mathbf{r}) + \underbrace{v_{xc}(\mathbf{r})\psi_i(\mathbf{r})}_{\text{Approximations}} = \epsilon_i(\mathbf{r})\psi_i(\mathbf{r}) \tag{3.11}$$

which is similar to Eq. 3.9, with Hartree potential, kinetic energy and local external potential terms appear in the same way, but the local KS potential replaced by the non-local $\int \Sigma(\mathbf{r}, \mathbf{r}'; \epsilon_i)\psi_i(\mathbf{r})d^3\mathbf{r}'$. It is worth noting that the wave functions $\psi_i(\mathbf{r})$ and the energies in Eq. 3.9 should not be taken as single-particle quantities, as the self-energy contains all the dynamical many-particle processes. Another difference to KS-DFT is that the interacting G obtained from Σ can directly producing the complete excitation spectrum, while KS results yields a series of single-particle states, and their corresponding eigenvalues in principle shall not be interpreted as excitation energies. Additionally, due to the non-locality of the Fock exchange, it cancels with the self-interaction contained in the Hartree potential. Nevertheless, the under-braced portion

of both the KS equation and the quasi-particle is where approximations come into play. One of the most well known way of finding approximations to the self-energy Σ is via the GW approximation.

3.1.2 The self-consistent GW Approximation

The GW approximation scheme has been derive and discussed extensively in but not limited to the following works:*Hedin* (1965); *Ehrenreich et al.* (1978); *Hedin* (1999); *Aryasetiawan and Gunnarsson* (1998b); *March* (1999). It has gained successes in the past for various solid-state band gaps and spectral properties calculations (*Aryasetiawan and Gunnarsson*, 1998c; *Godby et al.*, 1988), with mostly being G_0W_0, GW_0 or semi self-consistent GW. Solid GW calculations carried out in a fully self-consistent manner are being rare due to multiple reasons including computational challenges for large realistic systems, as well as the existing debate regarding the many-body self-consistency effect (*Ku and Eguiluz*, 2002; *Delaney et al.*, 2004).

However, scGW can also be appealing for it provides a consistent framework for treating the ground- and excited-states of the system. This advantage arises from the nonperturbative nature of scGW, as opposed to the perturbative G_0W_0. Moreover, in contrast to the perturbative theories, scGW produces ground and excited states that are independent of the starting point. Additionally, scGW is completely free of empirical parameters. This can be beneficial especially in the circumstances where the guidance from experimental data are limited. In this section, we present a fully self-consistent finite-temperature GW approximation implementation scheme in atomic orbital (AO) basis for periodic systems.

In the self-consistent GW approximation (scGW), G represent the one-body Green's function and W is defined as the dynamically screened Coulomb potential, as oppose to the bare Coulombs interaction in the exchange self-energy. This work uses the Matsubara formalism of the the Green's function, in which a temperature-

dependent feature is naturally embedded. The self-energy of GW approximation can be decomposed into two part, a static Hartree-Fock self-energy and a frequency dependent self-consistent correction term:

$$(\Sigma^{GW})^{\mathbf{k}}(i\omega_n) = (\Sigma_{\infty}^{GW})^{\mathbf{k}} + (\tilde{\Sigma}^{GW})^{\mathbf{k}}(i\omega_n). \tag{3.12}$$

, where $\omega_n = (2n+1)/\beta$ is the fermionic Matsubara frequency, β is the inverse temperature.

The dynamic part of Eq. 3.12 in time domain takes a simple yet elegant mathematical form $\Sigma = -GW$:

$$(\tilde{\Sigma}^{GW})^{\mathbf{k}}_{i\sigma,j\sigma'}(\tau) = -\frac{1}{N_k}\sum_{\mathbf{q}}\sum_{a,b} G^{\mathbf{k}-\mathbf{q}}_{a\sigma,b\sigma'}(\tau)\tilde{W}^{\mathbf{k},\mathbf{k}-\mathbf{q},\mathbf{k}-\mathbf{q},\mathbf{k}}_{i,a,b,j}(\tau) \tag{3.13}$$

with $\tilde{W}$ be the spin independent effective dynamical screened interaction.

In practice of scGW, the starting Green's function G_0 is initialized with either a HF or a DFT result. We use the non-orthogonal Guassian-based Bloch waves $\phi_i^{\mathbf{k}}(\mathbf{r})$ as the single-particle basis for periodic systems:

$$\phi_i^{\mathbf{k}}(\mathbf{r}) = \sum_{\mathbf{R}} \phi_i^{\mathbf{R}}(\mathbf{r}) e^{i\mathbf{k}\cdot\mathbf{R}} \tag{3.14}$$

, where $\mathbf{k}$ is the wave vector of the first Brillouin zone, $\mathbf{R}$ is the unit cell index and i is the Gaussian atomic orbital index. The summation in Eq. 3.14 allows the basis to extend over the lattice. We define the four component Coulomb interaction in the $\phi_i^{\mathbf{k}}(\mathbf{r})$ basis as:

$$U^{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4}_{i,j,k,l} = \int_{\Omega} d\mathbf{r}_1 \int_{\mathbb{R}^3} d\mathbf{r}_2 \phi_i^{\mathbf{k}_1^*}(\mathbf{r}_1)\phi_j^{\mathbf{k}_2}(\mathbf{r}_1)\frac{1}{|\mathbf{r}_1-\mathbf{r}_2|}\phi_k^{\mathbf{k}_3^*}(\mathbf{r}_2)\phi_l^{\mathbf{k}_4}(\mathbf{r}_2) \tag{3.15}$$

To preserve the translational invariance, it is assumed that $\mathbf{k}_1 - \mathbf{k}_2 + \mathbf{k}_3 - \mathbf{k}_4 = 0$.

The expression in 3.15 can be decomposed for the sake of storage into three- and two-body tensors with introducing another set of auxiliary Gaussian Bloch wave set:

$$U_{i,j,k,l}^{\mathbf{k}_1,\mathbf{k}_2,\mathbf{k}_3,\mathbf{k}_4} = \sum_{Q} V_{i,j}^{\mathbf{k}_1,\mathbf{k}_2}(Q) V_{k,l}^{\mathbf{k}_3,\mathbf{k}_4}(Q) \tag{3.16}$$

, where Q is the auxiliary index and $V_{i,j}^{\mathbf{k},\mathbf{k}'}(Q)$ is a three-point integral.

Following from Eq. 3.16, the effective screened interaction $\tilde{W}$ in Eq. 3.13 can then be expressed:

$$\tilde{W}_{i,a,b,j}^{\mathbf{k},\mathbf{k}-\mathbf{q},\mathbf{k}-\mathbf{q},\mathbf{k}}(\tau) = \sum_{Q,Q'} V_{i,a}^{\mathbf{k},\mathbf{k}-\mathbf{q}}(Q) \tilde{P}^{\mathbf{q}}(\tau; Q, Q') V_{b,j}^{\mathbf{k}-\mathbf{q},\mathbf{k}}(Q') \tag{3.17}$$

with

$$\tilde{P}^{\mathbf{q}}(\tau; Q, Q') = \frac{1}{\beta} \sum_{n} \tilde{P}_{QQ'}^{\mathbf{q}}(i\Omega_n) e^{-i\Omega_n \tau} \tag{3.18}$$

, where the auxiliary re-normalized polarization function $\tilde{P}^{\mathbf{q}}(i\Omega_n)$ forms a $N_Q \times N_Q$ matrix for every momentum $\mathbf{q}$:

$$\tilde{P}^{\mathbf{q}}(i\Omega_n) = [\mathbb{I} - \tilde{P}_0^{\mathbf{q}}(i\Omega_n)]^{-1} \tilde{P}_0^{\mathbf{q}}(i\Omega_n) \tag{3.19}$$

and

$$\begin{aligned} \tilde{P}_{0;Q,Q'}^{\mathbf{q}}(i\Omega_n) = \frac{-1}{N_k} \sum_{\mathbf{k}} \sum_{\sigma,\sigma'} \sum_{a,b,c,d} & V_{d,a}^{\mathbf{k},\mathbf{k}+\mathbf{q}}(Q) \\ & \times G_{c\sigma',d\sigma}^{\mathbf{k}}(-\tau) G_{a\sigma,b\sigma'}^{\mathbf{k}+\mathbf{q}}(-\tau) V_{b,c}^{\mathbf{k}+\mathbf{q},\mathbf{k}}(Q') \end{aligned} \tag{3.20}$$

, $\Omega_n = 2n\pi/\beta$ $(n = 0, \pm 1, \cdots)$ is the bosonic Matsubara frequency.

We refer the reader to references *Iskakov et al.* (2020a); *Yeh et al.* (2022) for a detailed discussion.

Self-consistency is achieved by computing Σ^{GW} and updating G by solving the Dyson equation 3.8 repeatedly. Compact in size but accuracy preserved imaginary

time and frequency grid are realized through orthogonal polynomial transformations and interpolations.

The total energy of a many-electron system is a explicit functional of the Green's function, and can be expressed in multiple ways. Examples includes the Luttinger-Ward functional and the Klein functional that are stationary at self-consistent G, and the Galitskii-Migdal formulation that is known to be nonvariational. For scGW that obeys the Dyson equation 3.8, all different approaches in principle yields the same solution to the total energy. The Galitskii-Migdal equation therefore appears to be appealing due to its simplicity:

$$E_{GM} = \frac{1}{2}\mathrm{Tr}[(h+F)\rho] + \frac{2}{\beta}\sum_{n}^{N_\omega} Re(\mathrm{Tr}[G(i\omega_n)\Sigma(i\omega_n)]) \tag{3.21}$$

The GW approximation has achieved great success for systems ranging from atoms, molecues to solids (*Aryasetiawan and Gunnarsson*, 1998b; *Caruso et al.*, 2013). However, it also suffers from certain drawbacks including but not limited to electron self-screening and wrong atomic limit. The latter one directly leads to its failure in the capturing of strong electronic correlations.

3.2 The Self-energy Embedding Theory

To properly address the correlation effects beyond GW, embedding theories such as density functional theory and dynamical mean-field theory (DFT+DMFT) (*Georges and Kotliar*, 1992; *Georges et al.*, 1996; *Kotliar et al.*, 2006), GW approximation and dynamical mean-Field theory(GW+DMFT) (*Biermann et al.*, 2003; *Werner and Casula*, 2016), density matrix embedding theory (DMET) (*Knizia and Chan*, 2012, 2013; *Welborn et al.*, 2016), G W and extended dynamical mean-field(GW+EDMFT) (*Boehnke et al.*, 2016; *Nilsson et al.*, 2017; *Petocchi et al.*, 2020). the self-energy embedding theory (SEET) (*Lan et al.*, 2016; *Zgid and Gull*, 2017a; *Lan and Zgid*,

2017; *Lan et al.*, 2017a; *Tran et al.*, 2018; *Rusakov et al.*, 2018; *Iskakov et al.*, 2020c; *Yeh et al.*, 2021) has been devised.

Among the above methods, SEET combines the weakly correlated GW approximation with non-perturbative quantum impurity solver, and has been tested extensively on 2D Hubbard model and molecular systems. In SEET, solutions to the theory are obtained in a thermodynamic consistent fashion, and conservation laws are satisfied. Unlike the DMFT methods that suffers from either the need of screened interaction terms or 'double-counting' and the impurity interaction parameters, SEET is free from any empirical parameter inputs.

We describe the fundamental setup of SEET in this chapter. It's application to realistic solids has been demonstrated in Chapter V on a particular transition metal oxide BiVO3.

3.2.1 SEET with GW

In the application of self-embedding theory to periodic systems, symmetrized atomic orbitals (SAO) are used for the construction of the electronic structure Hamiltonian. The fully self-consistent GW result $(G^{GW})^{\mathbf{k}}$ and $(\Sigma^{GW})^{\mathbf{k}}$ serves as a starting point for SEET. The former term is the GW momemtum-resolved Green's funtion of the lattice system, and is defined by the following Dyson equation:

$$(G^{GW}(i\omega_n))^{\mathbf{k}} = [(i\omega_n + \mu)\mathbb{I} - h^{0,\mathbf{k}} - (\Sigma^{GW}(i\omega_n))^{\mathbf{k}}]^{-1}. \tag{3.22}$$

The fundamental idea of self-energy embedding is to insert non-perturbative corrections to the GW self-energy Σ^{GW} on a selected subset of orbitals that could potentially have strong correlations involved. The self-energy expression in SEET is hence

modified to be:

$$(\Sigma^{SEET})^{\mathbf{k}} = (\Sigma^{GW})^{\mathbf{k}} + \sum_{A}((\Sigma_A^{imp})_{ij} - (\Sigma_A^{DC-GW})_{ij})_{\delta_{(i,j)\in A}}. \tag{3.23}$$

In Eq.3.23, Σ^{GW} represent the weakly correlated self-energy of the entire system, A is a selected subset of orbitals denoted as impurity, with indices $i, j \in A$. These so-called active orbitals or impurity orbitals are typically the most physically relevant orbitals for the correlated problem that requires treatment from quantum impurity solvers like Exact Diagonalization (ED). The term Σ_A^{imp} contains the non-perturbatively self-energy contribution in the orbital sets A, and the corresponding double counting correction is subtracted via the term Σ_A^{DC-GW}.

Consequently, the interacting propagator G of the lattice is defined as:

$$(G^{SEET}(i\omega_n))^{\mathbf{k}} = [(i\omega_n + \mu)\mathbb{I} - h^{0,\mathbf{k}} - (\Sigma^{SEET}(i\omega_n))^{\mathbf{k}}]^{-1}. \tag{3.24}$$

3.2.2 SEET Impurity Problem

To select the set of impurity orbitals local to each individual unit cell of the lattice, one needs to perform Fourier transform of $(G^{SEET}(i\omega_n))^{\mathbf{k}}$, $(\Sigma^{SEET})^{\mathbf{k}}$ and $h^{0,\mathbf{k}}$ from $\mathbf{k}$-space to real space, with Eq. 3.23 reformulated as:

$$(\Sigma^{SEET})^{\mathbf{RR'}} = (\Sigma^{GW})^{\mathbf{RR'}} + \sum_{A}((\Sigma_A^{imp})_{ij} - (\Sigma_A^{DC-GW})_{ij})_{\delta_{\mathbf{RR'}}\delta_{(i,j)\in A}} \tag{3.25}$$

, where $\mathbf{R}$ and $\mathbf{R'}$ are indices of unit cells. The term $\delta_{\mathbf{RR'}}$ implies that the non-perturbative self-energy correlation is applied to only the central unit cell of the lattice for selected impurity orbitals. This leads to the definition of Green's function restricted to subspace A in SEET:

$$(G^{SEET}(i\omega_n))_{ij\in A}^{\mathbf{RR}} = [(i\omega_n + \mu)\mathbb{I} - h_{ij\in A}^{0,\mathbf{RR}} - (\Sigma^{SEET})_{ij\in A}^{\mathbf{RR}} - \Delta_{ij\in A}^{A}(i\omega_n)]^{-1}. \tag{3.26}$$

, where Δ_{ij}^{A} is a matrix that contains the hybridization information between the impurity orbitals and the environment. It is therefore important to notice that $[(G^{SEET})_{ij\in A}^{\mathbf{RR}}]^{-1} \neq (G^{SEET,-1})_{ij\in A}^{\mathbf{RR}}$, and the differences between these two term is where the hybridization function stems from. Eq. 3.26 can be re-ordered into To solve for the for $\Sigma_A^{imp})_{ij}$, the auxiliary impurity $\mathcal{G}_A$ propagator is defined as follows:

$$\mathcal{G}_A^{-1} = (i\omega_n + \mu)\delta_{ij} - (h_{ij\in A}^{0,\mathbf{RR}} + (\Sigma_\infty^{GW})_{ij\in A}^{\mathbf{RR}} - (\Sigma_{\infty,A}^{DC})_{ij\in A}) - \Sigma_{ij\in A}^{imp} - \Delta_{ij\in A}^{A} \tag{3.27}$$

, with Σ_∞ be the static frequency independent self-energy.

3.2.3 SEET Self-Consistency

With the solved impurity problems and a known Σ_A^{imp}, one can update/evaluate the total self-energy and the corresponding Green's function of the lattice through Eq. 3.23 and 3.24. New self-consistent loop start by Fourier transform into real space (Eq. 3.26 and followed by a new solution to the impurity.

A self-consistency reached in Eq. 3.26 is viewed as a 'inner self-consistency loop' of SEET. The loop can be continued with an updated one iteration of GW solution followed by one iteration of ED impurity solver, and the ultimate converged solution can be obtained with this so-called 'outer self-consistency loop'.

CHAPTER IV

Gaussian Basis Sets for Solid-State Calculations

The work presented in this chapter has been published in the following paper:
Yanbing Zhou, Emanuel Gull, and Dominika Zgid, "Material-Specific Optimization of Gaussian Basis Sets against Plane Wave Data", *Journal of Chemical Theory and Computation* **2021** 17 (9), 5611-5622

4.1 Introduction

In simulations of molecular systems, choosing a basis set for representing the electronic wave function is a well established procedure (*Nagy and Jensen*, 2017; *Helgaker et al.*, 2000b; *Davidson and Feller*, 1986; *Huzinaga*, 1985). Gaussian basis sets that express a single orbital as a linear combination of primitive Gaussian orbitals are the most common molecular choice and multiple basis sets are available for most elements (*Schuchardt et al.*, 2007; *Pritchard et al.*, 2019; *Nagy and Jensen*, 2017; *Boys*, 1950).

In simulations of periodic solids, many more choices of computational bases are commonly employed, including plane waves (*Ihm et al.*, 1979; *Kresse and Furthmüller*, 1996a) and linear augmented plane waves (LAPW) (*Blaha et al.*, 2020b). Over the years, DFT (*Hohenberg and Kohn*, 1964a; *Kohn and Sham*, 1965) calculations of solids in the plane wave basis (*Ihm et al.*, 1979; *Kresse and Furthmüller*, 1996a) were

far more commonly used than Gaussian (*Martin*, 2004; *Dovesi et al.*, 2018) or any other basis set. This is due to the sparsity of the 2-body integrals in the plane wave representation, as well as the systematic convergence of the basis with respect to the number of plane waves (*Kresse and Furthmüller*, 1996b).

However, plane wave calculations are only affordable due to the low computational scaling of DFT. Thousands of plane waves are necessary to reach an accuracy comparable to that of a Gaussian basis set (*Kresse and Furthmüller*, 1996b; *Tosoni et al.*, 2007) with many fewer basis functions. Consequently, when considering post-DFT, ab-initio calculations with a higher scaling than the $\mathcal{O}(n^3)$ scaling of DFT in the number n of basis functions, Gaussian basis sets with fewer basis functions are an appealing choice. For instance, even simple post-DFT methods such as the second order Møller-Plesset perturbation theory (MP2) (*Ayala et al.*, 2001; *Pisani et al.*, 2005; *Maschio et al.*, 2007; *Usvyat et al.*, 2007; *Maschio et al.*, 2010; *Del Ben et al.*, 2012; *Usvyat et al.*, 2011) result usually in higher a computational scaling with respect to the number of basis functions than DFT.

Moreover, when a physical or chemical interpretation of results is desired, Gaussian basis functions provide an invaluable tool to gain direct insight into the behavior of electrons in chemically relevant atomic orbitals. While similar insight can be obtained in a plane wave or LAPW calculation, it requires an additional projection procedure, for example to Wannier orbitals (*Marzari et al.*, 2012; *Blaha et al.*, 2020a). Finally, any calculation that requires the inclusion of core orbitals, such as the study of core ionization processes, requires an explicit representation of those orbitals as provided by Gaussian basis sets, rather than an effective treatment via pseudopotentials.

While there are many advantages of employing a Gaussian basis for the description of solids, there are also clear drawbacks. First, while there is a systematic way of improving a plane wave basis by adding additional orthogonal high-energy functions, this is not easily possible in a Gaussian basis, since the basis functions do not form

an orthonormal set. It is frequently observed that it is impractical or impossible to reach convergence with respect to the number of basis functions by simply adding additional Gaussian orbitals. This is due to linear dependencies that appear once a large number of Gaussians with small exponents (describing diffuse orbitals) are present (*Cox and Fry*, 1977; *Karpfen*, 1979; *Suhai et al.*, 1982; *Kudin and Scuseria*, 2000; *Heyd et al.*, 2005; *Usvyat et al.*, 2011). Second, unlike in the case of plane waves, the exponents and coefficients of Gaussians need to be optimized for a given reference atomic problem. A minimization of the electronic energy with respect to a DFT calculation may optimize the basis function spanning the space of occupied orbitals (which determine the ground state energy) well, while leaving the ones necessary to express unoccupied orbitals relatively unoptimized.

Direct comparisons of results obtained with plane waves to those obtained with Gaussians are difficult, since plane wave bases rely on pseudopotentials, while Gaussian basis either explicitly treat core orbitals or employ different types of pseudopotentials. The differing description of core orbitals results in overall energy shifts of the remaining orbitals, making it somewhat cumbersome to compare the band structures between plane wave and Gaussian codes. Whether a differing band structure is a result of an insufficient Gaussian basis or of a different description of the core is often difficult to determine.

Despite of these disadvantages, the compactness of the basis and the ease of chemical/physical interpretation in the language of atomic orbitals are powerful advantages. However, the overall availability of Gaussian basis sets for solids is very limited, and the number of published basis sets for solids is a small fraction of the basis sets available for molecular systems (*Peintinger et al.*, 2013; *Vilela Oliveira et al.*, 2019; *Laun et al.*, 2018; *Daga et al.*, 2020; *Morales and Malone*, 2020). This is especially true for the 5th row transition metals such as lanthanides, and actinides that are common in many newly synthesized interesting compounds. It motivates our work on revisiting

the basis set optimization for solids.

One of the difficulties in choosing a good Gaussian basis for the description of solids is that the same element may display vastly different physical and chemical properties depending on its surroundings. For instance, different arrangements of carbon atoms create graphite, diamond, and graphene crystal lattices with both insulating and metallic character, respectively. Capturing these different behaviors with a single generic basis set, as is done in molecular systems, is difficult. We therefore explore a different paradigm in this work: rather than generating generic basis sets for solids, we optimize our basis sets in a material specific way (see also (*Daga et al.*, 2020) for related ideas).

The ground state energy remains an important target for the optimization. In addition, we minimize deviations from the orbital eigenvalues of a reference plane-wave calculation, while keeping basis functions as linearly independent as possible. Using the orbital eigenvalues of a reference calculation as an optimization target has two effects. First, it avoids getting trapped in local energy minima which occur in the optimization of deep-lying orbitals. Second, it allows us to also optimize the unoccupied orbitals, which do not directly enter the expressions for the ground state energy.

The material specific optimization approach retains the obvious advantage of Gaussian orbitals in that it produces compact bases with straightforward physical/chemical interpretation that can then be used in beyond-DFT calculations with higher-scaling methods. At the same time, by estimating the discrepancies between a Gaussian basis set and a non-Gaussian reference calculation, it eliminates the main disadvantage of working in a potentially incomplete basis set, and allows us to quantify basis set errors. We stress that this procedure is general and other reference calculations such as numerical bases or LAPWs may be used instead of plane waves.

In the reminder of this chapter, we describe the optimization algorithm (Sec. 4.2),

discuss results for commonly studied solids (Sec. 4.3), and present conclusions in Sec. 4.4.

4.2 Method

Our basis set construction consists of two parts: **(I)** *an atom-specific but materials-independent* generation of exponents and coefficients for an initial Gaussian basis set, followed by **(II)** the optimization of the initial basis set in *a material specific* manner for a given compound. In this section, we introduce Gaussian basis functions in 4.2.1, describe the initial optimization of the atomic problem in 4.2.2, and describe the subsequent material-specific optimization of the valence states in 4.2.3. The procedure is illustrated in Fig. 4.1.

4.2.1 Gaussian Orbitals

A basis set is defined as a set of single-particle functions (usually non-orthogonal atomic orbitals) that are then employed to build molecular orbitals (*Slater*, 1930). In Gaussian basis sets, contracted Gaussian-type orbitals (cGTOs) are used to approximate atomic orbitals. The expansion of a single cGTO employs a linear combination of primitive Gaussian-type orbitals (GTOs) (*Jensen*, 2006; *Helgaker et al.*, 2000a)

$$\Phi^{cGTO}(x,y,z) = N \sum_{i=1}^{N_{\text{prim}}} c_i \Phi_i^{GTO}(x,y,z), \tag{4.1}$$

where c_i is the contraction coefficient, N_{prim} describes the number of primitive GTOs, and N is a normalization constant. A primitive GTO, $\Phi^{GTO}(x,y,z)$ is expressed in terms of spatial coordinates x, y, and z and integer exponents a, b, and c as

$$\Phi^{GTO}(x,y,z) = \tilde{N} x^a y^b z^c e^{-\zeta r^2}, \tag{4.2}$$

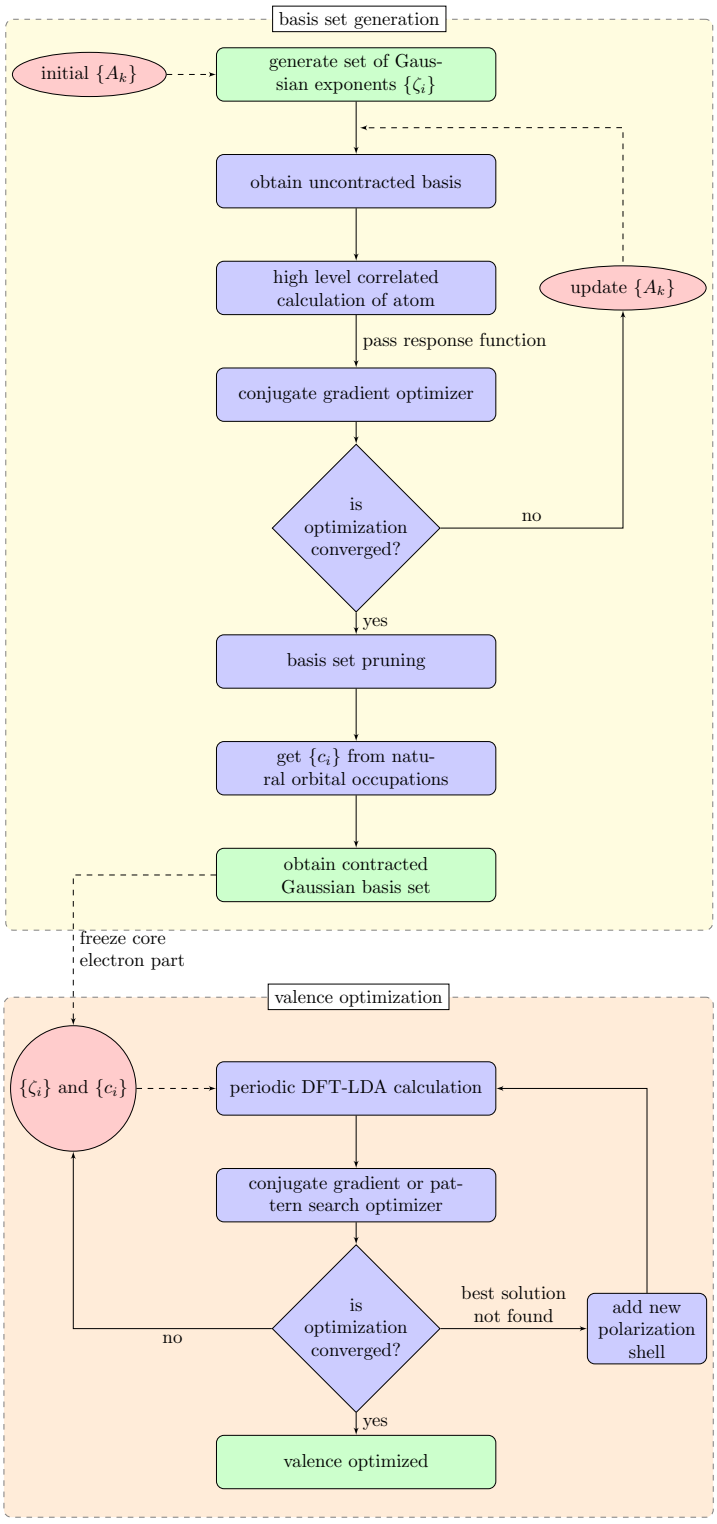


Figure 4.1: The basis set optimization scheme. Upper block: algorithmic steps for the generation of the initial exponents and contraction coefficients of the basis set. Lower block: steps for material-specific optimization of the valence orbitals.

where $a+b+c$ controls the angular momentum quantum number, ζ is the exponent of the GTO controlling the width of the orbital, and $\tilde{N}$ is a normalization constant. Large ζ results in a tight function while small ζ gives a diffuse function.

4.2.2 Generation of Initial Basis Set

The first step of our procedure (top panel of Fig. 4.1) generates exponents and contraction coefficients for each of the atoms. This procedure is not material specific and it is primarily used to optimize the cGTOs for the atomic core.

While the basis set generation procedure detailed above is not new, we describe it here for the sake of completeness. For many elements, it may be possible to start from existing atomic basis sets. However, for "more exotic" elements such as rare earth metals a generation of the atomic basis may be either necessary or desired.

4.2.2.1 Generation of starting exponents

To find an initial set of exponents ζ_i for the primitive Gaussians Φ_i^{GTO} that will later be assembled into cGTOs, we follow the procedure described in Ref. (*Petersson et al.*, 2003). There, the natural logarithm of each of the exponents ζ_i, $i = 1, \ldots, N_{\text{prim}}$ is expanded into an orthonormal Legendre polynomial expansion

$$\ln \zeta_i = \sum_{k=0}^{k_{\max}} A_k P_k\left(\frac{2i-2}{N_{\text{prim}}-1} - 1\right), \tag{4.3}$$

in terms of coefficients A_k which are common for different ζ_i exponents. N_{prim} is the number of Gaussian primitives in a given contracted orbital, P_k is the Legendre polynomial of order k. The initial set of parameters A_k are obtained according to the procedure described in Ref. (*Petersson et al.*, 2003). In our work, we set $k_{\max} = N_{\text{prim}} - 1$ or $k_{\max} = 6$, whichever is greater. The expansion into a Legendre basis eliminates the problem of linear dependencies that occurs when initial values of the

exponents are optimized independently.

4.2.2.2 Initial optimization of exponents and coefficients

With the initial set of $\{\zeta_i\}$ obtained, a first uncontracted atomic basis set is generated with a total number of $k_{max} + 1$ variational parameters A_k. This uncontracted basis is employed as a starting guess for the energy minimization in an atomic Hartree-Fock (HF) calculation. We use a conjugate gradient minimization (*Hestenes and Stiefel*, 1952; *Adams et al.*, 2014) scheme to find the variational parameters $\{A_k\}$ that minimize the electronic energy of the atom. Subsequently, the atom parameters are re-optimized in the configuration interaction singles and doubles (CISD) method.

Once the optimization of the ζ_i converges, contraction coefficients c_i for cGTOs can be found. They are generated as described in Ref. (*Almlof and Taylor*, 1987) by evaluating the expansion coefficients of the natural orbitals obtained by CISD in the basis of primitive GTOs. After this step, the number of cGTOs is then equal to the number of uncontracted functions.

4.2.2.3 Basis set pruning

Based on the natural orbital occupations of the CISD solution, many cGTOs typically have occupancies close to zero. We eliminate these cGTOs from the calculation by choosing to retain only the ones with an occupation threshold greater than 0.001.

In addition, diffuse Gaussian primitives with exponents smaller than 0.1 are eliminated, in analogy to Refs. (*Dovesi et al.*, 1990; *Peintinger et al.*, 2013; *Vilela Oliveira et al.*, 2019). This approximation was found to remove numerical instabilities in the calculation of solids. The resulting atomic orbitals can then be separated into core and valence orbitals.

4.2.3 Valence Optimization

The set of exponents and contraction coefficients obtained from Sec. 4.2.2.2 and pruned in Sec. 4.2.2.3 form the starting point of the valence optimization. Unlike the calculations in Sec. 4.2.2.2 and Sec. 4.2.2.3, this procedure is material specific and adjusts the valence orbitals to the chemical environment present in a given solid, while the core configuration is assumed to be material-independent and remains frozen.

A basis optimization where only the total energy E_{tot} is minimized has the tendency to optimize only the lowest eigenvalues well. However, the optimization of higher occupied orbitals may then become more difficult, since they contribute only little to the lowering of the overall DFT total electronic energy which is partially proportional to the sum of occupied orbital eigenvalues. As a consequence, the optimization may get stuck in a local minimum.

To remedy this issue, we choose an additional optimization criterion based on the difference between the molecular orbital (MO) energy at each k-point between the Gaussian calculation and a reference calculation. In our case, the reference is derived from a plane-wave calculation of the same system, taking advantage of the fact that the quality of the plane-wave basis set can be systematically controlled by a single parameter, the cut-off energy (e_{cut}), and can be converged with respect to the basis set size. Note, however that different pseudopotential present in the Gaussian and/or the plane wave calculation may lead to an overall shift of all energy bands. This shift is captured by an overall orbital-independent but potentially k-dependent energy shift parameter $\Lambda(k)$. The procedure of finding $\Lambda(k)$ is explained in Sec. 4.2.3.1.

The minimization of the deviation from a reference band structure is characterized by a functional $\|G(\{\zeta_i, c_i\}) - P(e_{cut}) - \Lambda(k)\|_F$, where $G(\{\zeta_i, c_i\})$ are MO eigenvalues of the Gaussian basis set calculation at each k-point, $P(e_{cut})$ are the corresponding plane-wave eigenvalues, and $\Lambda(k)$ is a k-dependent shift value. $||\cdot||_F$ denotes the Frobenius norm (*Golub and Van Loan*, 1996).

A third optimization criterion is given by the condition numbers (*Cheney and Kincaid*, 2007) of the overlap matrices at each k-point, which are minimized to obtain as little linear dependence in the resulting basis as possible. In a nearly linearly dependent Gaussian basis, the condition number can be very large or even infinite. Acceptable conditions numbers usually are smaller than 10^5. Consequently, denoting the vector of condition numbers of the overlap matrices as $\vec{\kappa}$, the minimization of the term $\gamma_2 \|\vec{\kappa}(\{\zeta_i, c_i\}\|_2)$, with $\|\cdot\|_2$ denoting the Euclidian norm, provides a penalty for linearly dependent solutions. Constraints on the condition number of the overlap matrix are frequently used in this context, see *e.g.* Refs. (*VandeVondele and Hutter*, 2007b; *Kuhne et al.*, 2020; *Lu et al.*, 2019; *Daga et al.*, 2020).

These three conditions lead to a minimization problem of the valence orbitals for a set of exponents of the Gaussian primitives and corresponding contraction coefficients $\{\zeta_i, c_i\}$ that minimize the functional

$$\begin{aligned} \Omega(\{\zeta_i, c_i\}) &= E_{tot}(\{\zeta_i, c_i\}) + \\ &+ \gamma_1 \|G(\{\zeta_i, c_i\}) - P(e_{cut}) - \Lambda(k)\|_F \\ &+ \gamma_2 \|\vec{\kappa}(\{\zeta_i, c_i\})\|_2, \end{aligned} \tag{4.4}$$

where γ_1 and γ_2 are weight factors that control the relative contributions of each of the three terms. In this work, γ_1 was set between 100 and 10^3 and γ_2 between 10^{-5} and 1.

In practice, we minimize the functional $\Omega(\{\zeta_i, c_i\})$ with a combination of a conjugate gradient optimizer and a non-gradient based pattern search (*Hooke and Jeeves*, 1961) optimization which we use to overcome local minima.

The bottom half of Fig. 4.1 illustrates the workflow for optimizing the basis functions for the valence orbitals. After convergence of the valence optimization, the core orbitals can be selectively re-optimized.

Note that the valence optimization method described here can also be applied to established basis sets, adapting a general-purpose basis set to a specific material context.

4.2.3.1 The value of Λ in the eigenvalue optimization

A systematic shift $\Lambda(k)$ is frequently present between the plane wave and Gaussian basis set codes. This quantity frequently makes the comparisons of band diagrams between the plane wave codes and Gaussian orbital codes inconvenient. While in general the value of Λ is k-dependent, in practice since the core orbitals are well localized they form flat and approximately k-independent energy bands. This is why in this work, we assume that Λ is k-independent.

The value of $\Lambda(k)$ can be evaluated by noticing several facts. First, we assume that the pseudopotential in the converged plane wave calculations was created to account for the low lying occupied orbitals and it recovers the energies of low lying occupied orbitals very accurately. (*Cao and Dolg*, 2011; *Dolg and Cao*, 2012) Then an all-electron calculation or a calculation with an accurate pseudopotential in an infinite/optimal Gaussian basis set can yield eigenvalues that are different only by a constant factor that is result of a freedom to add a constant shift in the pseudopotentials employed during the plane wave calculation.

Consequently, for every k-point, in an optimized bases the value of the shift Λ can be found using the least square fit between eigenvalues p_j in the plane wave basis and g_j Gaussian basis $\min \sum_j^{N_{eig}} (p_j - g_j - \Lambda)^2$, where N_{eig} is a chosen number of eigenvalues in the summation (N_{eig} can be equal to the number of occupied eigenvalues or it can also include some unoccupied eigenvalues, see the following discussion). We can find the shift Λ by finding the minimum of the above equation as a function of

Λ, $\frac{\partial}{\partial \Lambda}(\sum_j^{N_{eig}} (p_j - g_j - \Lambda)^2) = 0$. Finally, the value of the Λ shift is given by

$$\Lambda = \sum_j^{N_{eig}} \frac{(p_j - g_j)}{2N_{eig}}. \tag{4.5}$$

The knowledge of the value of Λ allows us to use the constraint: $\|G(\{\zeta_i, c_i\}) - P(e_{cut}) - \Lambda(k)\|_F$ helping the optimization of eigenvalues and to match approximately (within the desired criteria) the occupied eigenvalues. It also can be used to optimize and match a selected number of unoccupied eigenvalues. In this way, the quality of the Gaussian basis is rather systematically improved since we can choose how many unoccupied eigenvalues are being optimized. We can also observe which groups of AO orbitals have to be added to the basis to match increasingly higher lying virtual eigenvalues.

Moreover, finding the value of the shift described above, allows us to easily examine the quality of Gaussian basis sets prior to their usage by comparing the results to a converged plane wave calculation. Note also the even if a user does not desire to use such a specific constraint relying on matching the eigenvalues, it can be used to drive the optimization in the initial optimization steps and subsequently can be released allowing us to only insist on the optimization of the value of the total energy during the later steps.

4.3 Results

In this section, we assess the generation and optimization scheme described in the previous sections. Our discussion consists of two parts. In Sec. 4.3.1, we showcase the accuracy of the bases generated for simple monoatomic solids such as diamond, graphite, and silicon. We then use the optimization scheme developed in Sec. 4.2.3 to optimize the existing basis sets pob-DZVP-rev2(*Peintinger et al.*, 2013; *Vilela Oliveira et al.*, 2019; *Laun et al.*, 2018) and pob-TZVP-rev2 in a material specific way for the

solids BN, MnO, MoS_2, NiO. We also to assess the quality of existing GTH basis sets (*VandeVondele and Hutter*, 2007a).

Unless otherwise mentioned, all the calculations presented here using Gaussian basis sets are performed with CRYSTAL17 (*Dovesi et al.*, 2017) using DFT with the LDA. The reference calculations were performed in a plane wave basis in GPAW (*Mortensen et al.*, 2005) with an energy cutoff of 2000 eV and a Monkhorst-Pack grid of $8 \times 8 \times 8$ k-points. We use the DAKOTA (*Adams et al.*, 2014) conmin-frcg DAKOTA coliny-pattern-search package as optimizers. For all cases analyzed, the optimizer converges within 100 optimization steps.

4.3.1 Basis generation for diamond, graphite, and silicon

This basis generation scheme is intended to be material specific and a different basis set should be generated for every compound examined. While some of these bases for a given atom may be transferable between different materials, we generally believe that due to significantly different properties of solids containing the same atoms (eg. diamond, graphite, and graphene), a materials specific basis optimization is preferred in calculations of solids.

The initial atomic basis sets for C and Si are generated using the method described in Sec. 4.2.2.1 and schematically presented in the yellow shaded block of Fig. 4.1. Note that the high level method used to perform correlated calculations in an uncontracted basis for C and Si atoms is configuration interaction singles and doubles (CISD) from Psi4 (*Parrish et al.*, 2017). This step allows us to find the initial set of exponents and contraction coefficients that are then pruned, taking advantage of natural orbitals. Subsequently, each shell present in the core part of the basis is first optimized individually using the functional from Eq. 4.4. The resulting core basis functions are then frozen and the valence basis functions are optimized by the conjugate gradient optimizer, with a pattern search optimizer as a further refinement tool. This opti-

Basis Sets	Diamond	Graphite	Silicon
OPT	4s3p1d	6s5p1d	4s3p1d
CATTI-1993	3s2p1d	3s2p1d	*
DOVESI-1990	3s2p1d	3s2p1d	*
VALENZANO-2006	3s2p1d	3s2p1d	*
GATTI-1994	3s2p1d	3s2p1d	*
NADA-1996	*	*	4s3p1d
PASCALE-2005	*	*	5s4p1d
NADA-1990	*	*	4s3p1d
DARCO-1993	*	*	5s4p1d
TOWLER-1998	*	*	6s5p1d
POB-TZVP-2012	4s3p1d	4s3p1d	5s4p1d
HEYD-2005	4s3p1d	4s3p1d	6s5p1d
POB-DZVP-REV2	3s2p1d	3s2p1d	4s3p1d
POB-TZVP-REV2	4s3p1d	4s3p1d	5s4p1d
dcm[C_{diam}]-TZVP	5s3p2d1f	*	*

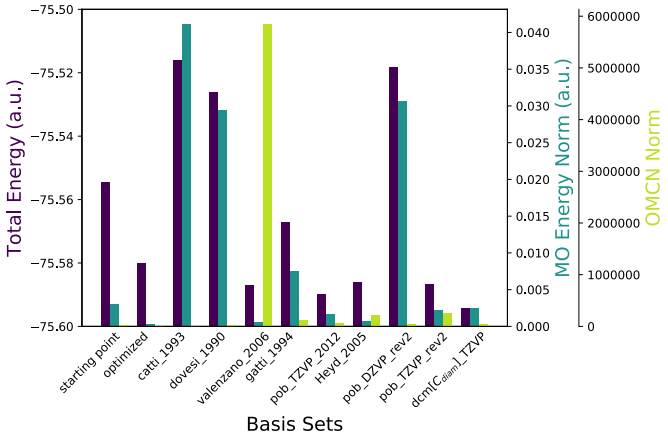
Table 4.1: Number of shells in each of the basis sets listed in Fig. 4.2. A star (*) denotes that a given basis was not employed in a given simulation.

mization procedure can be done for the atom or in a material specific way, where the valence part of the basis is then reoptimized again for each of the compounds.

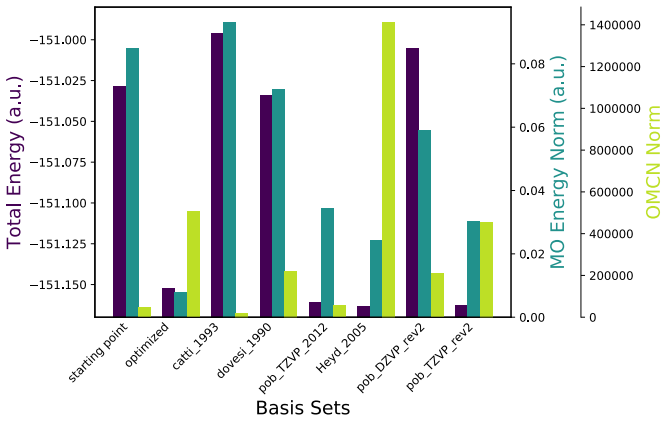
To examine the quality of the generated basis sets, we compare them to a list of existing basis sets for solid state calculations (*CRYSTAL*, 2021) of diamond, graphite, and silicon. In Fig. 4.2 we show the results of the comparison between the basis set obtained from the atomic level optimization, the material specific optimized basis set, and the existing basis sets listed in literature.

For diamond (see Fig. 4.2 panel a)) the basis set generated has 4 s shells, 3 p shells, and 1 d shell. The number of primitives in each shell is illustrated in Tab. 4.3.

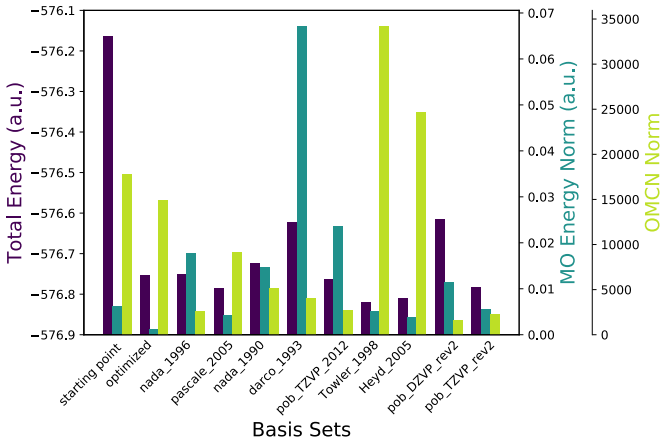
We observe that this initial atomic basis set, before the material specific optimiza-



(a)



(b)



(c)

Figure 4.2: A comparison of generated bases with existing Gaussian bases for :(a) diamond;(b) graphite;(c) silicon. On the Y axis, Total Energy stands for the total energy obtained in an LDA calculation $E_{tot}(\{\zeta_i, c_i\})$ from Eq. 4.4, MO Energy Norm corresponds to the $\|G(\{\zeta_i, c_i\}) - P(e_{cut}) - \Lambda(k)\|_F$ term from Eq. 4.4, OMCN stands for the overlap matrix condition number which is the 3rd term $\|\vec{\kappa}(\{\zeta_i, c_i\})\|_2$ from Eq. 4.4.

tion, has moderately high electronic energy comparable to older available basis sets such as the one optimized by Gatti (*Gatti et al.*, 1994) in 1994. However, both its energy and its MO norm are much lower than pob-DZVP-rev2 (*Vilela Oliveira et al.*, 2019). The optimized basis is comparable to pob-TZVP-rev2 (*Vilela Oliveira et al.*, 2019) and the original pob-TZVP (*Peintinger et al.*, 2013) basis as well as the basis developed by Heyd (*Heyd et al.*, 2005). While the electronic energy in these bases is slightly lower, the overlap norm and the MO eigenvalue norm are higher.

We also compared our optimized basis to the material specific optimized basis set for diamond listed in Ref. (*Daga et al.*, 2020), that we call it here dcm[C_{diam}]-TZVP. The electronic energy in the dcm[C_{diam}]-TZVP basis is somewhat lower than in pob-TZVP-rev2 while the overlap norm is greatly minimized. while our optimized basis yields a higher energy than the dcm[C_{diam}]-TZVP basis, we have significantly fewer basis functions in our basis explaining the difference in energy. For the number of basis functions in each of the listed bases see Tab. 4.1. These results suggest that both of these bases that were optimized in material specific way are good alternatives to the standard library basis sets and they can yield excellent energies while avoiding problems with the linear dependencies appearing in bases for periodic problems. A list of the most diffuse exponents present in the bases listed in Fig. 4.2 is given in Tab. 4.2.

For graphite (see Fig. 4.2 panel b)), the basis set generated has the same shell structure as the diamond basis, while the number of Gaussian primitives are slightly different, as shown in Tab. 4.4. The optimized basis is again comparable in its quality to the pob-TZVP-rev2 (*Vilela Oliveira et al.*, 2019). However, it has a much lower MO norm while the norm of the overlap remains comparable.

For silicon (see Fig. 4.2 panel c)), the basis set generated has 4 s shells, 3 p shells, and 1 d shell. The number of primitives in each shells is shown in Tab. 4.5. The optimized basis yields comparable results to the pob-TZVP bases. However, the

Basis Sets	Diamond	Graphite	Silicon
OPT	0.226	0.189	0.110
CATTI-1993	0.260	0.260	*
DOVESI-1990	0.196	0.196	*
VALENZANO-2006	0.146	0.146	*
GATTI-1994	0.169	0.169	*
NADA-1996	*	*	0.193
PASCALE-2005	*	*	0.130
NADA-1990	*	*	0.170
DARCO-1993	*	*	0.270
TOWLER-1998	*	*	0.127
POB-TZVP-2012	0.164	0.164	0.115
HEYD-2005	0.185	0.185	0.120
POB-DZVP-REV2	0.210	0.210	0.140
POB-TZVP-REV2	0.164	0.164	0.135
dcm[C_{diam}]-TZVP	0.201	*	*

Table 4.2: The exponent of the most diffuse GTO for each basis sets. A star (*) denotes that a given basis was not employed in a given simulation.

optimized basis has a lower MO norm while having a slightly higher overlap norm.

In general, the total energies of our basis sets are comparable to the pob-TZVP basis sets, while the MO energy norms are much lower and the overlap condition numbers remain reasonable. This is partly due to the fact that we are optimizing multiple parameters at once, i.e., optimizing both the eigenvalue norm and the total energy together, instead of minimizing the energy alone. Note that the MO constraint causes the equal optimization of all eigenvalues, independent of their magnitude. This means that we can design a basis to optimize a particular set of eigenvalues of interest, either in the occupied part, or in the unoccupied part. We can also choose to optimize a certain number of bands in a specific energy window.

It is also worth mentioning that, although the three terms in $\Omega(\{\zeta_i, c_i\})$ from Eq. 4.4 may appear independent of each other, mutual dependencies between them are observed. As demonstrated in Fig. 4.2, the MO norm and the total energy value varying patterns are consistent with each other independent of a specific basis set. When the MO norm is high the total energy associated with it also tends to be higher, and once the MO norm is minimized the resulting energy is low. This trend is expected and confirms that MO norm is a good indicator of the basis set quality and can be used as a driving factor during the basis set optimization. The overlap norm anti-correlates with the value of the total energy. When the overlap norm is very small, the total energy tends to increase. This is expected, as the optimization of any basis has to balance orthogonality of the basis functions with the overall the lowering of the total energy.

4.3.2 Material specific optimization of existing bases

4.3.2.1 MoS_2

For MoS_2 (Pearson symbol: hP6, crystal system: hexagonal, space group number: 194) the calculations performed in the pob-DZVP-rev2 basis show that the high

Shell Index	s-shell	p-shell	d-shell
1	6		
2	4	4	
3	1	1	1
4	1	1	

Table 4.3: Number of Gaussian primitives in each shell in the optimized C-diamond basis set.

Shell Index	s-shell	p-shell	d-shell
1	6		
2	4	3	
3	1	1	1
4	1	1	

Table 4.4: Number of Gaussian primitives in each shell in the optimized C-graphite basis set.

Shell Index	s-shell	p-shell	d-shell
1	8		
2	8	4	
3	1	2	1
4	1	1	
5	1	1	
6	1	1	

Table 4.5: Number of Gaussian primitives in each shell in the optimized Si basis set.

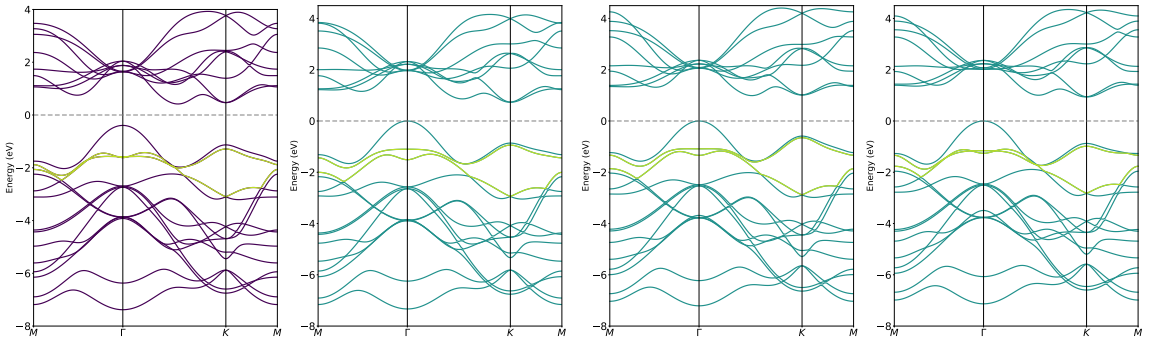


Figure 4.3: MoS_2 band structure plots: (a) plane-wave basis calculation with cut off energy: 2000eV; (b) Gaussian basis calculation with original pob-DZVP-rev2-basis; (c) Gaussian basis calculation with original pob-TZVP-rev2-basis; and, (d) Gaussian basis calculation with optimized pob-TZVP-rev2-basis.

lying occupied band marked in light green and shown in Fig. 4.3 panel b) do not show the band crossings that are present in the reference plane waves calculations. The calculations conducted in the pob-TZVP-rev2 presented in Fig. 4.3 panel c) do not show a qualitative improvement when compared to pob-DZVP-rev2 results. Optimization procedures following Sec. 4.2.2.3 are performed with Mo's ECP (*Andrae et al.*, 1990), 4s, 4p and 4d shell frozen and S's 1s, 2s and 2p core frozen. We observe that with the existing number of shells in the basis set, the optimization of pob-DZVP-rev2 basis does not recover the band crossing present in the reference. For the pob-TZVP-rev2 basis, our optimization scheme helps to improve the predicted band structure, see Fig. 4.3 panel d), and is showing a consistently better band structure behavior when comparing to the reference result. Fig. 4.4 shows the convergence of the material specific optimization process where we start with the pob-TZVP-rev2 basis. The values of the three individual terms of Eq. 4.4: the value of the total energy, the norm of the overlap matrix, and the difference in the MO eigenvalues for every k-point after accounting for the presence of the constant shift are displayed separately for each of the cycles.

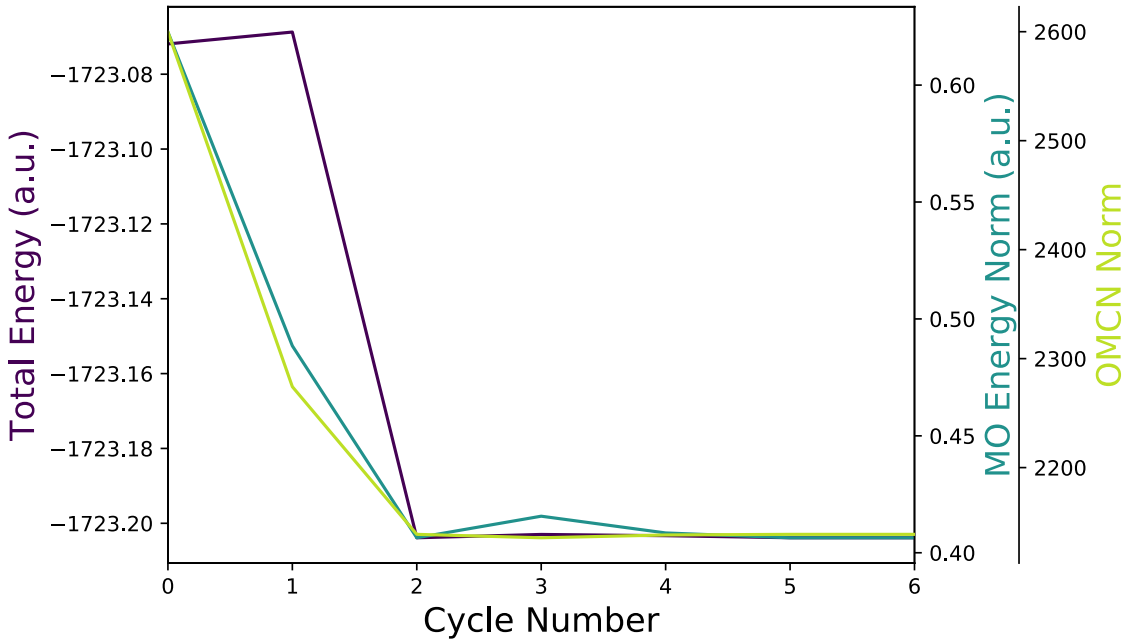


Figure 4.4: Values of the total energy (purple curve), MO energy norm (dark green curve), and overlap norm (light green curve) during the minimization of the functional Ω of Eq. 4.4 for solid MoS_2. The minimization was performed using a conjugate gradient optimizer. For details see Sec. 4.3.2.1.

In the optimization process, we find that the Mo optimized pob-TZVP-rev2 is significantly different from its library version. The changes due to the optimization of the pob-TZVP-rev2 basis of S are less pronounced but also important.

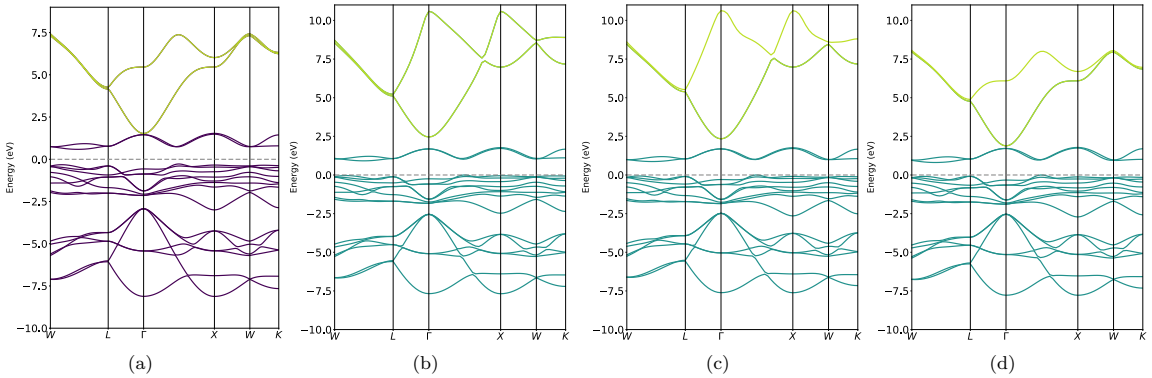


Figure 4.5: NiO band structure plots: (a) plane-wave basis calculation with cut off energy: 2000eV; (b) Gaussian basis calculation with original pob-DZVP-rev2-basis; (c) Gaussian basis calculation with original pob-TZVP-rev2-basis; and, (d) Gaussian basis calculation with pob-TZVP-rev2-basis and an additional d shell.

4.3.2.2 NiO

For NiO (Pearson symbol: cF8, crystal system: fcc rocksalt, space group number 225, AFM) the calculated band structure Fig. 4.5 panel b) and Fig. 4.5 panel c) using pob-DZVP-rev2 and pob-TZVP-rev2 basis,respectively are both qualitatively inconsistent with the reference Fig. 4.5 panel a) for the unoccupied band marked in light green. Moreover, the pob-TZVP-rev2 basis does not lead to any significant improvement over pob-DZVP-rev2. We attempted the optimization of both pob-DZVP-rev2 and pob-TZVP-rev2. However, an optimization with the existing number of shells in these bases did not improve the quality of the resulting band structure. To remedy this, we added an additional 4d valence shell to the oxygen atom in the pob-TZVP-rev2-basis. In Fig. 4.5 panel d), we illustrate that after optimizing the basis set in this updated format, the inconsistency for the unoccupied bands of the

Gaussian basis set is fully removed when compared to the plane wave basis.

4.3.3 Assessment of the basis set quality

The difference in the eigenvalues between the plane wave and Gaussian basis set $\|G - P - \Lambda\|_F$ can be used as a criterion for a quick assessment of the basis quality in addition to the visual inspection of the differences. Here, we examine differences between the pob-DZVP-rev2 and pob-TZVP-rev2 bases for the NiO and MoS_2 solids studied previously, in addition to hexagonal BN and MnO solids. We show the band diagrams below for BN and MnO. the band diagrams for NiO and MoS_2 were presented in the previous section.

4.3.3.1 BN

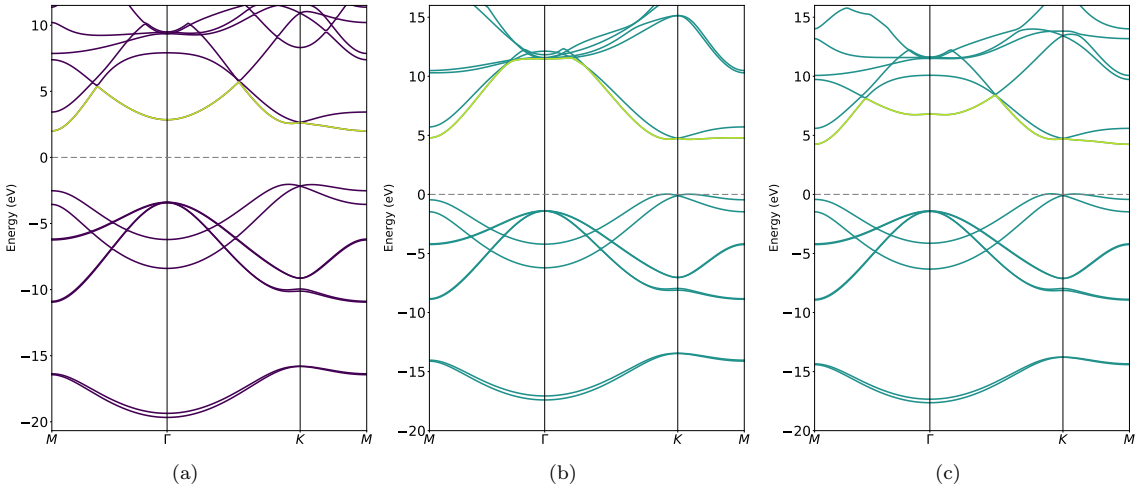


Figure 4.6: BN band structure plots: (a) plane-wave basis calculation with cut off energy: 2000eV; (b) Gaussian basis calculation with pob-DZVP-rev2-basis; (c) Gaussian basis calculation with pob-TZVP-rev2-basis.

For BN (Pearson symbol: hP4, crystal system: hexagonal, space group number: 194), as shown in Fig. 4.6, the pob-DZVP-rev2-basis shows a qualitative discrepancy in the unoccupied part of the band structure near the Fermi level (that is marked

in light green on the band structure diagram) when compared to the plane-wave reference calculations. The pob-TZVP-rev2-basis is qualitatively consistent with the reference.

4.3.3.2 MnO

For MnO (Pearson symbol: cF8, crystal system: fcc rocksalt, space group number: 225, AFM) the pob-DZVP-rev2 basis produces a better band structure result than the pob-TZVP-rev2-basis when compared with the plane wave reference. This indicates that the extra shells added in the TZVP basis are not essential or not optimized enough for a quantitatively correct description of the higher lying bands of MnO. This is illustrated in Fig. 4.7, where the incorrectly recovered bands are marked in light green.

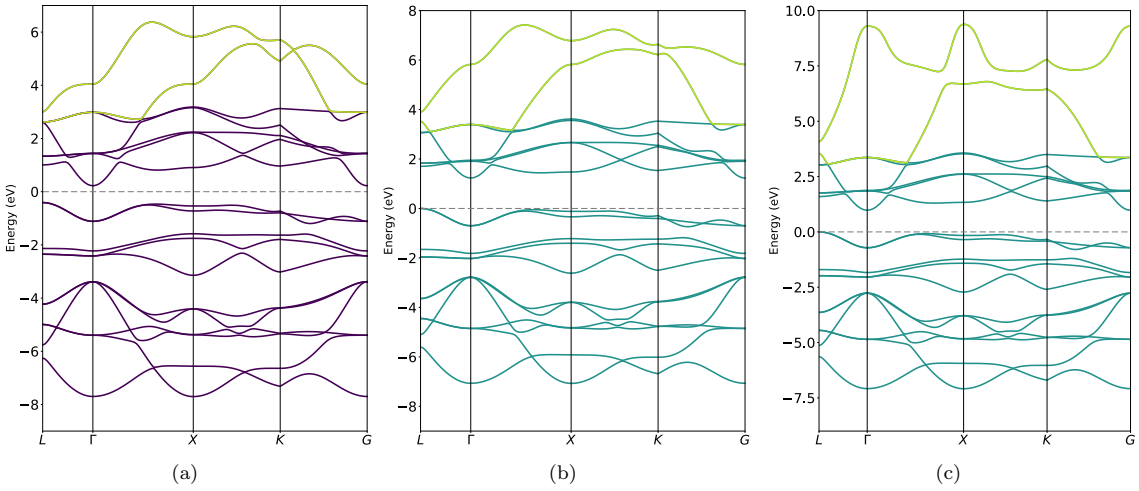


Figure 4.7: MnO band structure plots: (a) plane-wave basis calculation with cut off energy: 2000eV; (b) Gaussian basis calculation with pob-DZVP-rev2-basis; (c) Gaussian basis calculation with pob-TZVP-rev2-basis.

In Tab. 4.6, we show the MO eigenvalue norm evaluated for pob-DZVP-rev2 and pob-TZVP-rev2 for the systems studied in this paper. We stress that this norm is evaluated for the valence orbitals and does not include core orbitals. The norms for

the valence occupied and unoccupied orbitals are printed separately.

From Tab. 4.6, we see that for BN there is a dramatic improvement once the basis set is increased.

For NiO, the quality of the pob-DZVP-rev2 and pob-TZVP-rev2 is unchanged and is consistent with the band diagram plotted in Fig. 4.5. However, our observation does not indicate that the total DFT energy evaluated in pob-TZVP-rev2 is higher than in the pob-DZVP-rev2 basis. The pob-TZVP-rev2 energy is lower due to the optimization that happened for the core orbitals, while the valence bands remained mostly unaffected.

For MnO, pob-TZVP-rev2 is worse for the unoccupied bands than pob-DZVP-rev2, according to the assessment using the difference of eigenvalues. This is consistent with the band diagram plotted in Fig. 4.7.

For MoS_2, the quality of pob-DZVP-rev2 and pob-TZVP-rev2 seems to be similar with some small worsening of the unoccupied part in the pob-TZVP-rev2 basis.

Subsequently, we also employ this criterion to analyze briefly and schematically the quality of the GTH basis (*VandeVondele and Hutter*, 2007a) for BN. GTH bases are another family of the systematically developed bases that are commonly used for solid state calculation.

In the BN case, see Tab. 4.7, where both DZ and TZ bases are available, the gth-dzvp-molopt-sr basis with gth-pade as the pseudo-potential is yielding a similar deviation of the norm as the gth-tzvp-molopt-sr with the same pseudopotential. However, the maximal difference in the eigenvalues is larger for the unoccupied part of gth-tzvp-molopt-sr. This is consistent with the band diagram illustrated in Fig. 4.8, where a spike in the unoccupied part of the band diagram arises for the gth-tzvp-molopt-sr basis.

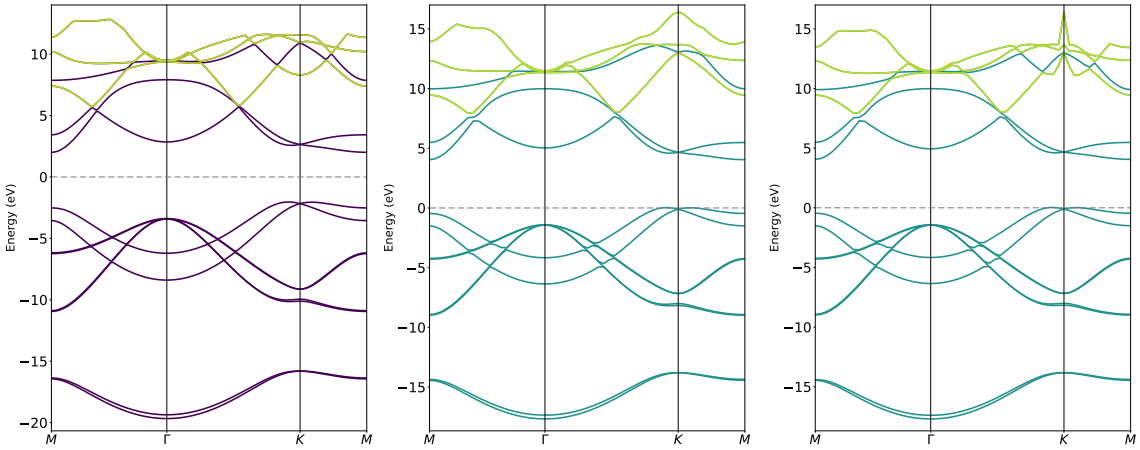


Figure 4.8: BN band structure plots calculated with GTH basis sets: (a) plane-wave basis calculation with cut off energy: 2000eV; (b) Gaussian basis calculation with DZVP-GTH basis; (c) Gaussian basis calculation with TZVP-GTHP basis.

System	OCC		UNOCC	
	DZVP	TZVP	DZVP	TZVP
BN	0.7186	0.2096	10.7039	4.2203
NiO	0.5911	0.7998	5.9098	5.9848
MnO	0.7328	0.7906	1.6521	2.3830
MoS_2	0.8715	0.8703	0.6222	0.7930

Table 4.6: MO eigenvalue norms calculated with respect to plane-wave basis calculations. Basis sets used are pob-DZVP-rev2 and pob-TZVP-rev2, respectively. BN: 8 occupied bands and 6 unoccupied bands are used; NiO: 22 occupied bands and 6 unoccupied bands are used; MoS_2: 10 occupied bands and 6 unoccupied bands are used; MnO: 11 occupied bands and 6 unoccupied bands are used.

Basis	Norm		Max. Diff	
	OCC	UNOCC	OCC	UNOCC
DZVP	0.2260	3.0873	2.0691	4.8523
TZVP	0.2432	3.0329	2.0668	4.8795

Table 4.7: BN MO eigenvalue norm calculated with respect to plane-wave basis calculations. Basis sets used are DZVP-GTH and TZVP-GTH, respectively, with pseudo-potential GTH-PADE. 8 occupied bands and 6 unoccupied bands are used in the norm calculation.

4.4 Conclusions

We have presented a Gaussian basis set optimization scheme for solid state calculations. This optimization scheme results in the minimization of the total DFT energy, the condition number of the overlap matrix, as well as minimization of the difference between the eigenvalues evaluated in the reference plane wave basis and the currently optimized Gaussian basis set.

Our aim of designing such a material specific optimization scheme is to employ the moderately sized optimized bases in correlated periodic *ab-initio* calculations such as second-order Green's function theory (GF2) (*Dahlen and van Leeuwen*, 2005; *Phillips and Zgid*, 2014; *Phillips et al.*, 2015; *Rusakov et al.*, 2014; *Rusakov and Zgid*, 2016; *Welden et al.*, 2016; *Kananenka et al.*, 2016a,b; *Kananenka and Zgid*, 2017; *Gull et al.*, 2018), GW (*Hedin*, 1965; *Aryasetiawan and Gunnarsson*, 1998a; *van Schilfgaarde et al.*, 2006; *Stan et al.*, 2006; *Koval et al.*, 2014; *Luo et al.*, 2002; *Caruso et al.*, 2016; *Holm and von Barth*, 1998; *Lan et al.*, 2017b), and self-energy embedding theory (SEET) (*Kananenka et al.*, 2015; *Zgid and Gull*, 2017b; *Rusakov et al.*, 2019; *Iskakov et al.*, 2020b) in order to evaluate and interpret spectral functions of solids. Gaussian bases are very useful for the physical interpretation of the peaks present in spectral functions and for assigning them to electron attachment or removal processes from s, p, or d orbitals. Since these *ab-initio* methods are computationally demanding,

a moderately sized Gaussian basis is extremely advantageous in carrying out these calculations and comparing them to experiment.

We believe that this scheme will be particularly useful for periodic calculations, where a given element may display vastly different properties depending on the surrounding crystal lattice environment. A significantly changed basis may then capture the changes in the valence orbitals arising from different physical/chemical properties of the crystal, while remaining of moderate size.

The improvement of basis sets quality for crystalline systems is not as straightforward as for molecular ones, for multiple reasons. We have shown in this work that for some compounds simply increasing basis set size by following the typical double, triple, and quadruple zeta hierarchy (eg. using a pob-TZVP-rev2 instead of pob-DZVP-rev2 for NiO) may not necessarily lead to an improvement of the basis set quality without additional material specific optimization. Moreover, the onset of linear dependence issues in periodic systems may be delayed by having a material specific optimization, thereby allowing one to employ bases of moderate sizes that are geared towards recovering experimental results. The advantage of the scheme presented here is the possibility of building a basis set of user desired quality by insisting that band after band is optimized when compared to a plane wave calculation. Alternatively, if so desired by a user, a series of bands within a given energy window can be optimized.

The obvious drawback of the scheme proposed here is a possible lack of transferability between different compounds. Moreover, in certain cases, such as energy assessment for different polymorphs, a different basis for each of the polymorphs would render such comparison invalid. However, even such a case, it is possible to optimize a single basis that is common for all the group of polymorphs and evaluate energy in such a basis. Our main motivation of designing this material specific scheme is to use it for *ab-initio* post-DFT calculations of strongly correlated compounds. For these

materials, few basis sets are readily available thus an optimization is frequently necessary before starting calculations. Moreover, for some of the materials of interest the experimental data may be scarce and calculations should have predictive quality. In such cases, material specific optimization is helpful to ensure that even a moderately sized basis set is the most optimal for the calculated material and all possible differences when comparing against experimental data come for the computational method employed. Finally, we believe that as post-DFT, correlated, *ab-initio* calculations for solids in Gaussian bases will become more prevalent, a database of material specific bases can be created, thus alleviating the issue of transferability.

At last, it is worth reiterating that the scheme presented here is not only useful for basis set optimization but also for a quick assessment of the basis set quality as illustrated in Sec. 4.3.3. In such a case, a score based on the eigenvalue differences between the plane wave and Gaussian basis can be assigned in order to aid the visual inspection necessary for the assessment of the basis set quality.

CHAPTER V

An ab-initio Tale of $BiVO_3$ Under Pressure

5.1 Introduction

Computational post-DFT studies of oxide perovskites received a lot of attention since they contain transition metal atoms that contribute to the onset of strongly correlated behavior in these compounds. Consequently, multiple methods such as LDA+DMFT or GW+EDMFT were used to study these compounds. Numerous previous studies were conducted not always fully ab-initio and have used adjustable parameters U and J to illustrate the emergence of strong correlation as a function of these parameters. Only recently fully ab-initio descriptions of oxides perovskites become available. Almost all such descriptions are done when the experimental data are present allowing the theorists to verify their predictions. In this work, it is our aim to conduct the description of $BiVO_3$ in a completely ab-initio manner without the full possibility of verification based on the experimental data. $BiVO_3$ is synthesized in a diamond anvil under pressure where only the X-ray crystalographical data are collected. Such synthesis was described in Ref. *Klein et al.* (2019). Consequently, it is impossible to assess if the synthesized compound is insulating or metallic based solely on the experimental data present.

Since perovskites are ideal candidates for the materials' design with specific magnetic and electronic functionalities, many of the newest perovskites based materials

lead to promising photovoltaics, photocatalysts, and multiferroics. Bismuth–first-row transition-metal oxide perovskites, $BiMO_3$, are particularly interesting (M = Sc, Cr, Mn, Fe, Co, Ni) due to a range of possible chemical/physical behaviors. In $BiMnO_3$, the lone pair of Bi electrons is allegedly giving rise to (anti)ferroelectric behavior, while the electronic spins on the Mn ions couple ferromagnetically, making $BiMnO_3$ a magnetoelectric multiferroic.

The family of $BiMO_3$ perovskites is almost completely synthesized and explored. Only the perovskites with M = Ti^{3+} and V^{3+} were not previously synthesized. Bulk, crystalline $BiVO_3$ remained synthetically inaccessible at ambient pressure because the Bi^{3+} ions oxidize the V^{3+} ions to form V^{5+} species and metallic bismuth Bi^0. In such a case, the high pressure synthesis enables the isolation of phases that, under ambient pressures, are inaccessible.

This is why the theoretical investigation of the magnetic and electronic properties of $BiVO_3$ is very important since it can shed some light on its properties without scaling up the synthesis route.

Previous calculations *Klein et al.* (2019) have shown that cubic BiVO3 is antiferromagnetic and metallic. This is an unusual case which deserves further computational investigations. Commonly, undoped, magnetic transition-metal oxides fall into two groups: semiconducting antiferromagnets and metallic ferromagnets. Only a couple of metallic antiferromagnetic transition metal oxides are known: $SrCrO_3$, $CaCrO_3$, $Ca_3Ru_2O_7$, $LaNiO_3$, $LaCu_3$ Cr_4O_{12}, RuO_2, and $(La,Sr)_3Mn_2O_7$. The exact mechanism for the anomalous antiferromagnetism in these metallic systems remains poorly understood. We belive that our ab-initio studies can shed some light onto these antiferromagnetic phases and investigate the possibility of band gap opening as the pressure is decreased after the synthesis in the diamond anvil.

5.2 Results

5.2.1 Computational Setup

We refer our readers to Chapter III and Appendix A for a comprehensive description of the theoretical setup of the self-consistent GW approximation and the self-energy embedding theory (SEET) employed by this implementation.

Computationally, to describe the A- and C-antiferromagnetic orderings we double the unit cell along the two different directions. The resulting unit cells are tetragonal (A-phase) and orthorombic (C-phase) with two vanadium metal atoms per cell to enforce antiferromagnatec orderings. We perform our calculations in a Gaussian gth-dzvp-molopt-sr basis with gth-pbe pseudopotential and decompose the four-fermion Coulomb integrals employing the density fitting in the def2-svpri as the auxiliary basis. We use up to $4 \times 4 \times 4$ k-points in the Brillouin zone for both GW and SEET calculations. All the integrals as well as DFT calculations are carried out with the open source PySCF package.

All quantities in SEET are computed on the imaginary time and frequency axis. In this work, we use a compact intermediate representation (IR) grid with sparse frequency sampling for all dynamical quantities such as Green's functions and self-energies. In IR, the grid size is governed by a dimensionless parameter Λ that should be at least larger than $\beta \times \omega_{max}$ where β is the inverse temperature and ω_{max} is the energy bandwidth of the system. Simulations are performed at temperature T= 1579K (β = 200 1/a.u.). The density of states (DOS) is computed based on a converged single-particle Green's function that is then analytically continued from the imaginary to the real frequency axis using the Nevanlinna analytical continuation along a high-symmetry k-path. This Nevanlinna continuation method ensures causality of the continued function. A broadening parameter of η = 0.001 a.u. is used for all calculations which sharpens the results. We find that such a broadening is crucial for

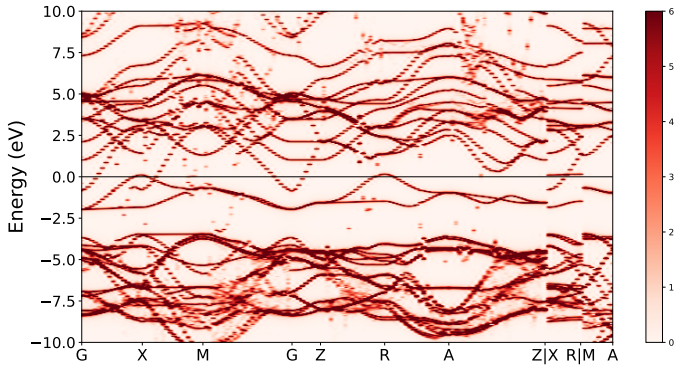
achieving a sharp band structure necessary for the discussion of the emerging band gaps.

The simplest possible strongly correlated subspace for $BiVO_3$ is made out of local vanadium 3d orbitals (split into t_{2g} and e_g). However, since $BiVO_3$ has nominally only a single electron in the vanadium d-shell, we consider only the t_{2g} impurity in its SEET description as is standard in many dynamical mean field theory (DMFT) calculations.

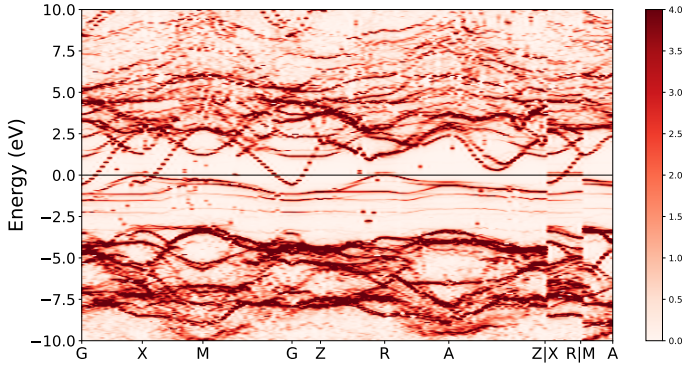
5.2.2 A-phase

In Fig. 5.1, we present results for the highest experimentally obtained pressure with a=3.791 Å in the unit cell. It is visible that the PBE0 functional yields a metallic solution due to the band crossing at the X point but also due to bands crossing the Fermi level at the Γ and R points. While the self-consistent, temperature dependent GW seems to be reducing the band crossing at the X point, overall the compound remains metallic. Similarly in SEET(GW/ED) the result is metallic. Both the GW and SEET results are in qualitative agreement with the result obtained from the PBE0.

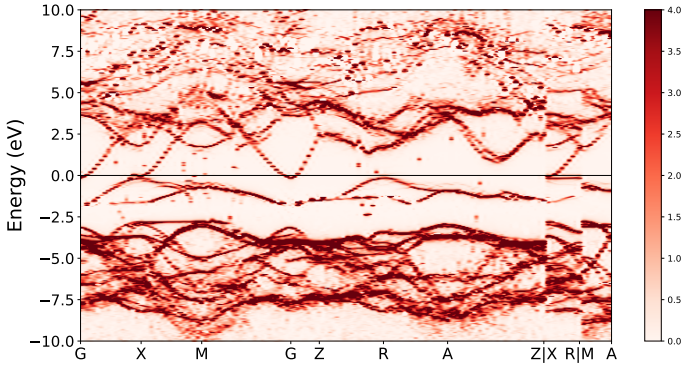
Similarly, in Fig. 5.2, for the lowest experimentally obtained pressure, we observe a metallic character for all the PBE0, GW, and SEET(GW/ED) solutions. Finally, we theoretically "depressurize" the A-phase even further to achieve the ambient pressure resulting in a=3.935 Å for the unit cell. In this case, we start to observe somewhat qualitative disagreement between the GW and PBE0 results. GW is barely metallic with only a single bands touching Fermi level at the X point while DFT shows crossing of multiple bands in the vicinity of X point. SEET(GW/ED) further opens a gap resulting in a semiconductor like physics and a narrow bandgap of less than 0.5 eV. In Fig. 5.4, we list set of orbitals that are energetically closest to the Fermi level. It is evident that Bi p_y and p_z as well as V d_{xy} are responsible for the band metallic or



(a)

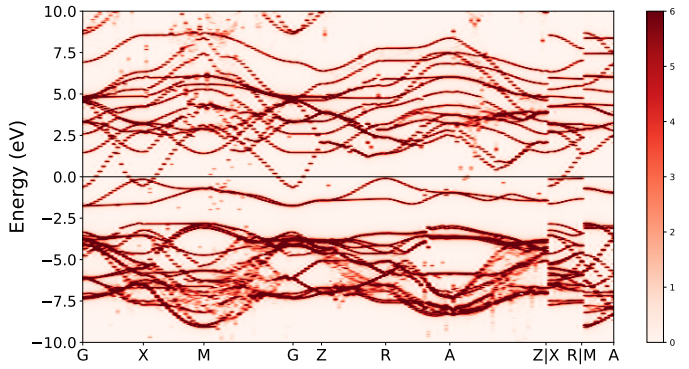


(b)

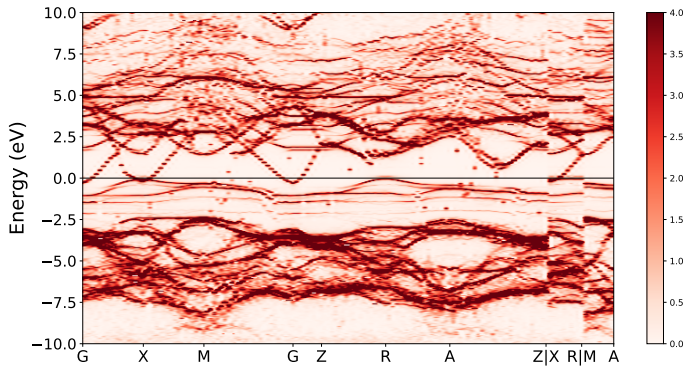


(c)

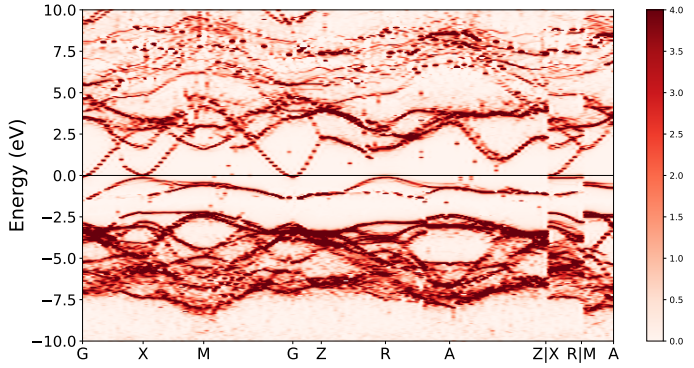
Figure 5.1: The band structure of the A-antiferromagnetic phase of $BiVO_3$ with unit cell distance a=3.791 Å corresponding to the highest experimentally reached pressure from Ref. *Klein et al.* (2019). *Top:* DFT band structure using the PBE0 functional *Middle:* Band structure using self-consistent finite temperature GW *Bottom:* Band structure using SEET(GW/ED)



(a)



(b)



(c)

Figure 5.2: The band structure of the A-antiferromagnetic phase of $BiVO_3$ with unit cell distance a=3.927 Å corresponding to the lowest experimentally reached pressure from Ref. *Klein et al.* (2019). *Top:* DFT band structure using the PBE0 functional *Middle:* Band structure using self-consistent finite temperature GW *Bottom:* Band structure using SEET(GW/ED)

semiconductor-like character in $BiVO_3$.

5.2.3 C-phase

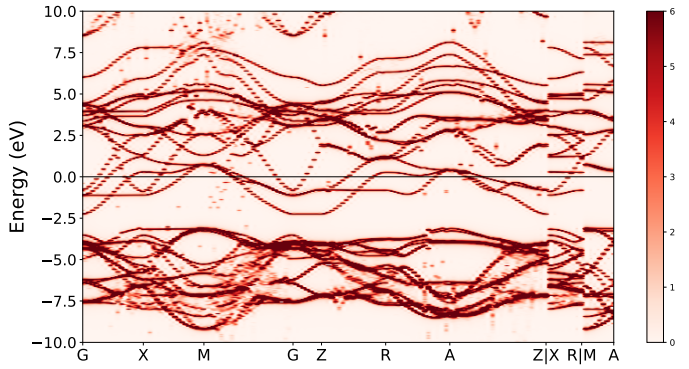
We conduct similar studies for the C-phase. Here, we start with analyzing the highest experimentally reached pressure in Fig. 5.5. Here PBE0 results in a band crossing along multiple k-points yielding a metallic solution. Both GW and SEET, also result in a metallic solution, however, here bands do not cross and the C-phase shows a semi-metallic character both at the GW and SEET level. The difference with the PBE0 behavior is rather stark. Even though PBE0 also predicts metallic solution, such a metal would have most likely quite different properties.

In both the lowest experimentally realized pressure and the lowest pressure to which we extrapolated the results, we show that both GW and SEET consistently predict semi-metal while PBE0 predicts a qualitatively different metallic solution. This behavior can be traced to the wrong illustration of the occupied V d_{yz} band that is too high in energy for PBE0, see Fig. 5.8.

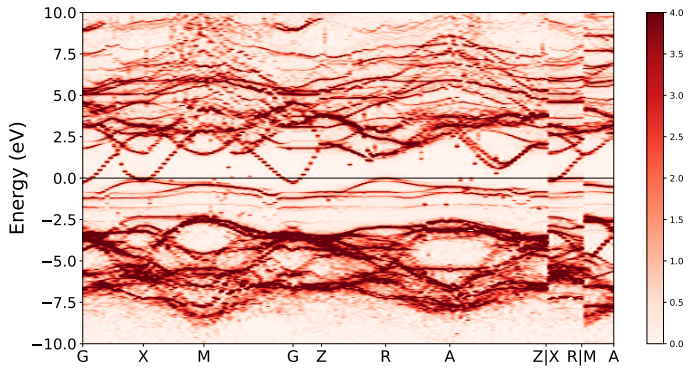
5.3 Conclusions

We have presented a band structure from post-DFT methods such as SEET for the newly synthesized $BiVO_3$ under pressure. The resulting band structure shows a failure of DFT in illustrating both insulating and metallic solutions. While for the A-phase in the high and lower pressure regimes, DFT gave quantitatively correct answer, it did not predict an insulating phase for the lowest extrapolated pressure. For the C-phase, DFT shows a qualitative breakdown yielding a metal instead of semi-metal.

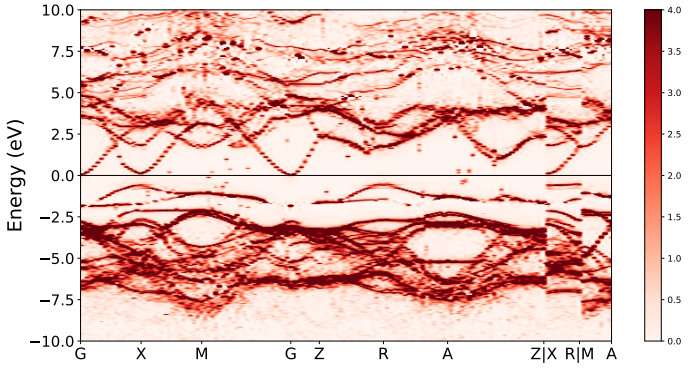
Our results demonstrate that the known dangers of using DFT when the experimental data may be very scarce. While this is not surprising, it should be noted that here it resulted in both quantitative and qualitative differences. The comparison



(a)



(b)



(c)

Figure 5.3: The band structure of the A-antiferromagnetic phase of $BiVO_3$ with unit cell distance a=3.935 Å corresponding to the lowest extrapolated pressure from Ref. *Klein et al.* (2019). *Top:* DFT band structure using the PBE0 functional *Middle:* Band structure using self-consistent finite temperature GW *Bottom:* Band structure using SEET(GW/ED)

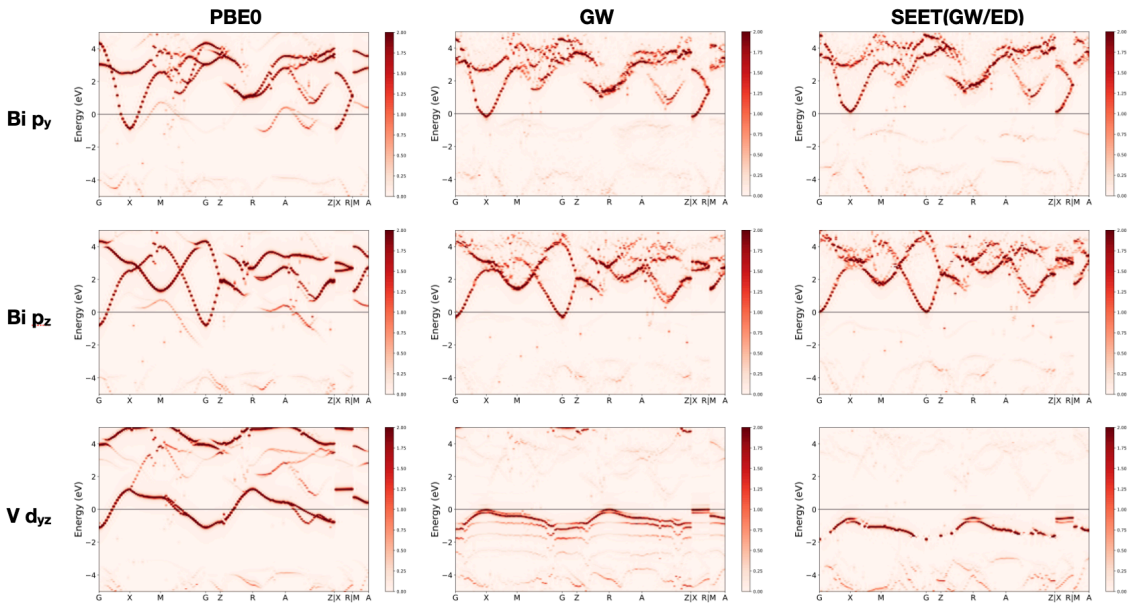
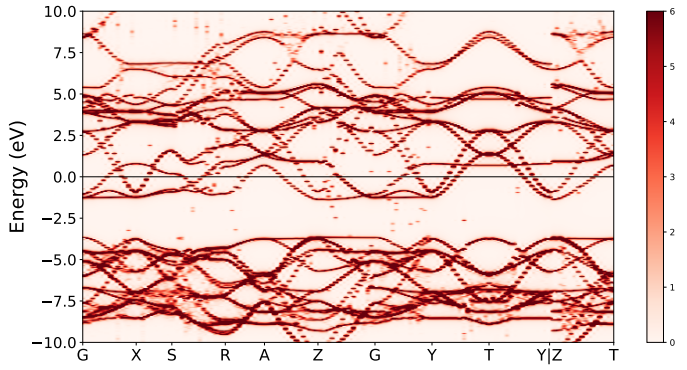
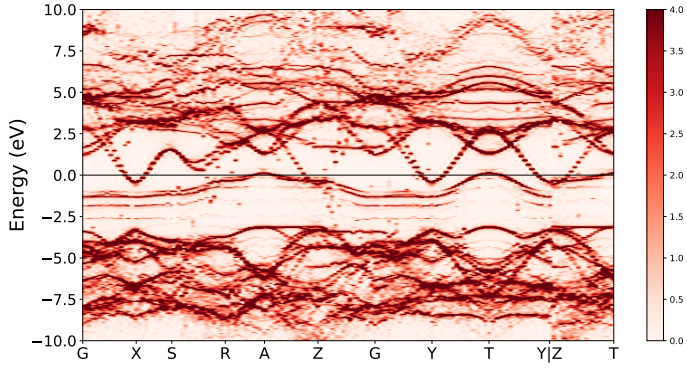


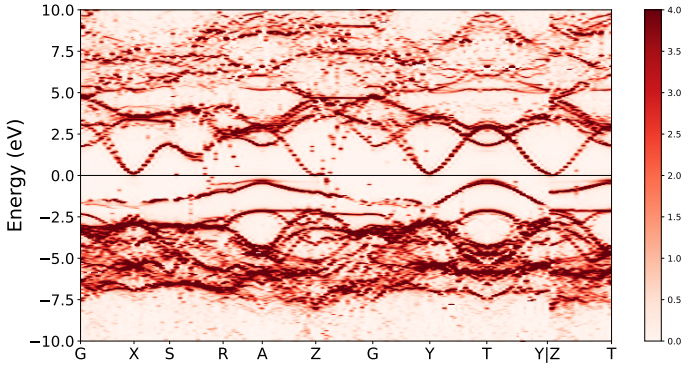
Figure 5.4: orbitals closest to the Fermi level for the A-antiferromagnetic phase of $BiVO_3$ with unit cell distance a=3.935 Å corresponding to the lowest extrapolated pressure from Ref. *Klein et al.* (2019). *Left:* DFT band structure using the PBE0 functional *Middle:* Band structure using self-consistent finite temperature GW *Right:* Band structure using SEET(GW/ED)



(a)

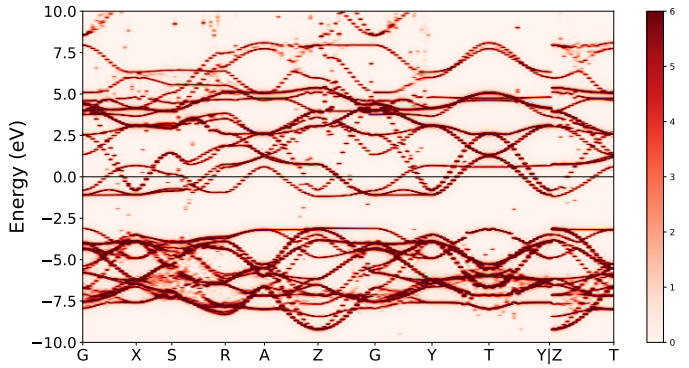


(b)

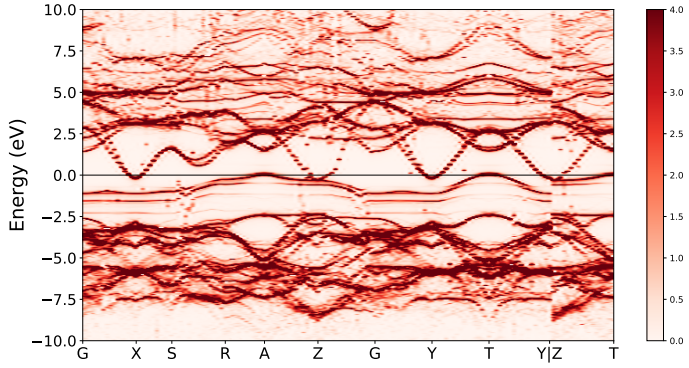


(c)

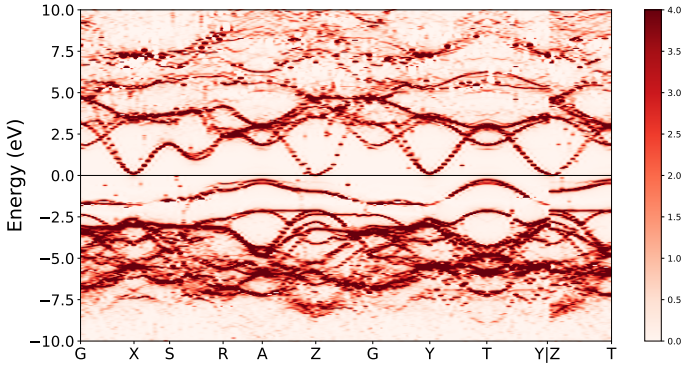
Figure 5.5: The band structure of the C-antiferromagnetic phase of $BiVO_3$ with unit cell distance a=3.791 Å corresponding to the highest experimentally reached pressure from Ref. *Klein et al.* (2019). *Top:* DFT band structure using the PBE0 functional *Middle:* Band structure using self-consistent finite temperature GW *Bottom:* Band structure using SEET(GW/ED)



(a)

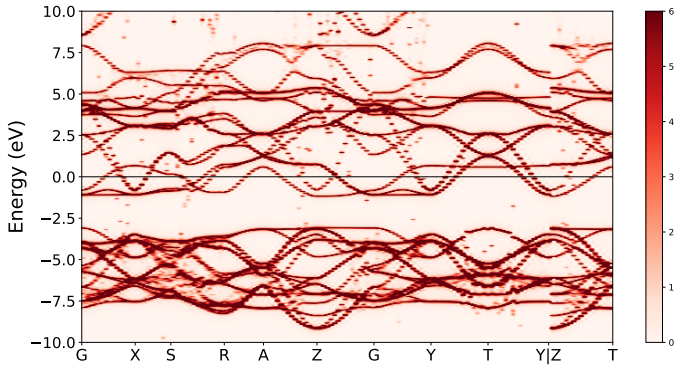


(b)

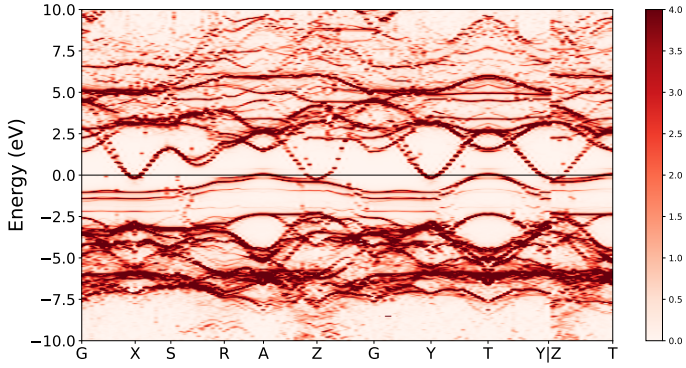


(c)

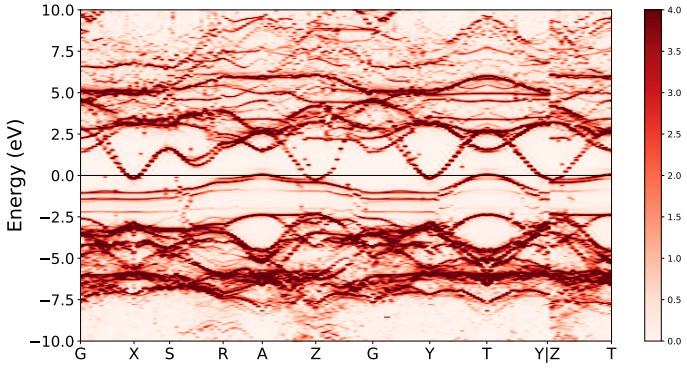
Figure 5.6: The band structure of the C-antiferromagnetic phase of $BiVO_3$ with unit cell distance a=3.927 Å corresponding to the lowest experimentally reached pressure from Ref. *Klein et al.* (2019). *Top:* DFT band structure using the PBE0 functional *Middle:* Band structure using self-consistent finite temperature GW *Bottom:* Band structure using SEET(GW/ED)



(a)



(b)



(c)

Figure 5.7: The band structure of the C-antiferromagnetic phase of $BiVO_3$ with unit cell distance a=3.935 Å corresponding to the lowest extrapolated pressure from Ref. *Klein et al.* (2019). *Top:* DFT band structure using the PBE0 functional *Middle:* Band structure using self-consistent finite temperature GW *Bottom:* Band structure using SEET(GW/ED)

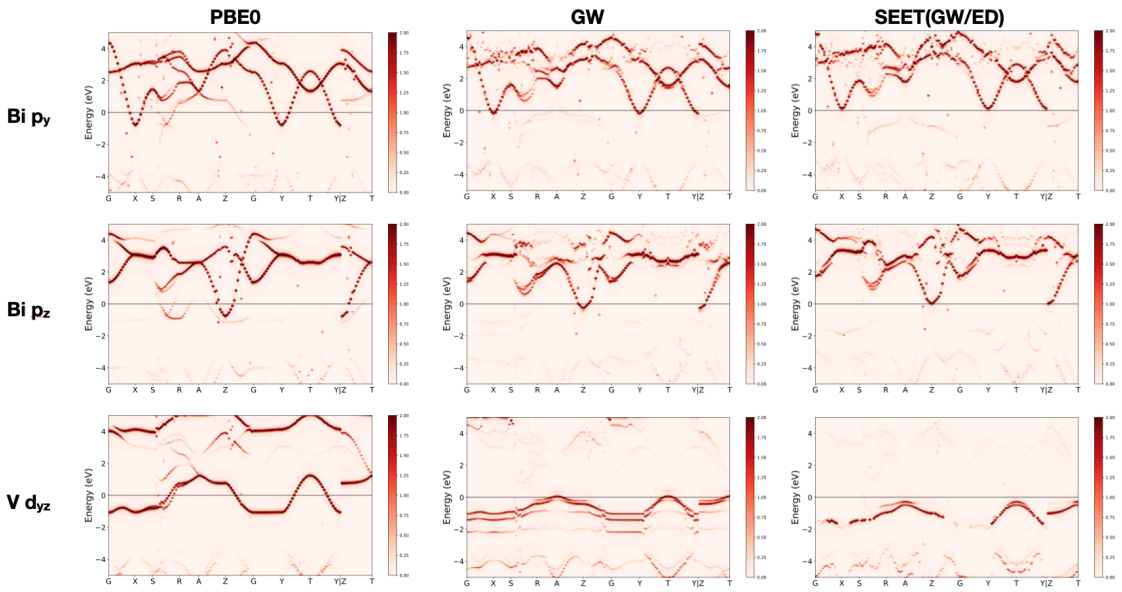


Figure 5.8: orbitals closest to the Fermi level for the C-antiferromagnetic phase of $BiVO_3$ with unit cell distance a=3.935 Å corresponding to the lowest extrapolated pressure from Ref. *Klein et al.* (2019). *Left:* DFT band structure using the PBE0 functional *Middle:* Band structure using self-consistent finite temperature GW *Right:* Band structure using SEET(GW/ED)

between GW and SEET data turned to be reassuring since both of them recover data that is both qualitatively and quantitatively similar.

CHAPTER VI

Conclusion

First principle computational simulations for the correlated electronic properties of materials could aid and guide the experimental synthesis and discovery of new correlated materials that could potentially benefit the advancement of modern-day technologies. The impediments to an efficient correlated simulation typically include the inappropriate choice of basis set used in the simulation, and failure in the capturing of strong electronic correlations. This thesis, however, has contributed to the solid-state quantum chemistry community by developing tools to circumvent the former hindrance and bench-marking ab-initio correlated methods designed to unravel the latter. The thesis was divided into two major topics: the study of designing and optimizing Gaussian basis sets for solid-state quantum simulations and the implementation of ab-initio self-consistent GW approximation and self-energy embedding theory (SEET) on a transition metal cubic perovskite, BiVO3.

6.1 Summary

In Chapter IV, we proposed a material-specific Gaussian basis sets optimization approach, which is shown to be computationally feasible and is tested to yield improved result against the existing solid-state Gaussian basis sets. We designed the proposed scheme to conduct correlated periodic ab-initio calculations such as GW

and SEET, for which good-quality but moderate-in-size basis sets are needed. This scheme has the feature of simultaneously minimizing the total energy of the system, the condition number of the overlap matrix, as well as minimizing the difference between the eigenvalues evaluated in the reference plane wave basis and the currently optimized Gaussian basis set. The basis quality improvement achieved by employing the proposed scheme has been demonstrated in the work. Additionally, using the scheme as a fast basis set quality assessment tool is recommended.

In Chapter V, We have analyzed the band structure properties of the cubic perovskites BiVO3 with correlated d electrons by systematically adding non-perturbative correlations to orbitals adjacent to Fermi energy. The results of BiVO3 simulation overall demonstrates an inconsistent prediction of metallic character between the DFT-PBE0 and our formulation. Individual band properties analysis exemplifies a quantitative, in some cases, even qualitative adjustment to the transition metal t_{2g} and e_g bands. This illustrates the failure of DFT in describing both insulating and metallic solutions of BiVO3 under various pressure and antiferromagnetic phases. And this was able to confirm the known dangers of using DFT when experimental data may be deficient. At the meantime, this work also reiterates that the Bloch-wave basis formed by Gaussian orbitals is handy for physically interpreting the individual bands from given atomic orbitals and assigning them to electron excitation processes.

6.2 Outlook

Moving forward, there are many directions in which the material-specific Gaussian basis optimization routine is improvable. As illustrated in 4.1, the basis generation and optimization procedure in this work is limited to a full electron formalism. Nevertheless, due to the extensive system size of the typical solid-state simulations, it is not rare for pseudo-potentials to be employed. Therefore, it is beneficial to have a tool that could assess and potentially generate good-quality yet computationally efficient

pseudo-potentials. Having a tool to aid in the quality assessment and generation of good-quality yet simulation-wise efficient peuso-potentials is therefore beneficial to the computationalist in the field.

Given the material-specific nature of the scheme, a distinct basis set might be required for each specific polymorph. We believe that a database of material-specific bases can be created to alleviate such cumbersome operations. The scheme could also be incorporated to provide benefit to current existing solid-state quantum chemistry packages as a preparatory optimization tool to improve the long-term simulation quality from inexpensive short-term input.

While the material-specific scheme is currently being implemented in the scope of DFT, the primary motivation for designing this scheme is for its application in ab initio post-DFT calculations of strongly correlated compounds. Its transferability can therefore be examined on selected computationally affordable systems. Moreover, with a growth of interest in solid-state relativistic scGW and SEET, the applicability of this scheme can be explored in the relativistic Gaussian basis sets.

APPENDICES

APPENDIX A

Ab-initio Self-consistent GW Approximation and Self-energy Embedding

S. Iskakov, C. Yeh, E. Gull, and D. Zgid

Physical Review B **102**, 085105 (2020)

We model a solid as an arrangement of atoms in a Bravais lattice with periodicity in all three directions. We employ the Born-Oppenheimer approximation and choose a basis of single-particle wave functions. In this work we use Bloch waves constructed from Gaussian basis functions as

$$\phi_{\mathbf{k}_i,i}(\mathbf{r}) = \sum_{\mathbf{R}} \phi_i^{\mathbf{R}}(\mathbf{R}) e^{i\mathbf{k}\cdot\mathbf{R}}, \tag{A.1}$$

where $\phi_i^{\mathbf{R}}(\mathbf{r})$ is a Gaussian atomic orbital centered in Bravais lattice cell $\mathbf{R}$. These states are not orthogonal and define the overlap matrix

$$\mathbf{s}_{ij} = \int_{\Omega} d\mathbf{r} \phi^*_{\mathbf{k}_i,i}(\mathbf{r}) \phi_{\mathbf{k}_j,j}(\mathbf{r}) \delta_{\mathbf{k}_i,\mathbf{k}_j}. \tag{A.2}$$

The electronic structure Hamiltonian in second quantization is

$$H = \sum_{\boldsymbol{ij},\sigma} h^0_{\boldsymbol{ij}} c^\dagger_{\boldsymbol{i}\sigma} c_{\boldsymbol{j}\sigma} + \frac{1}{2} \sum_{\substack{\boldsymbol{ijkl} \\ \sigma\sigma'}} v_{\boldsymbol{ijkl}} c^\dagger_{\boldsymbol{i}\sigma} c^\dagger_{\boldsymbol{k}\sigma'} c_{\boldsymbol{l}\sigma'} c_{\boldsymbol{j}\sigma}. \tag{A.3}$$

Where $c_{\boldsymbol{i}\sigma}$ ($c^\dagger_{\boldsymbol{i}\sigma}$) are annihilation (creation) operators corresponding to the single particle state $\phi_{\mathbf{k}_i,i}(\mathbf{r})$, with spin σ and index $\boldsymbol{i}(\boldsymbol{j},\boldsymbol{k},\boldsymbol{l})$ denotes the combined orbital-momenta index $\boldsymbol{i} = (i, \mathbf{k}_i)$. The single-particle operator $h^0_{\boldsymbol{ij}}$ and two-particle operator $v_{\boldsymbol{ijkl}}$ are defined respectively as

$$h^0_{\boldsymbol{ij}} = \int_\Omega d\mathbf{r} \phi^*_{\mathbf{k}_i,i}(\mathbf{r}) \left[-\frac{1}{2}\nabla^2_{\mathbf{r}} - \sum_\alpha \frac{Z_\alpha}{r_{\alpha,\mathbf{r}}} \right] \phi_{\mathbf{k}_j,j}(\mathbf{r}), \tag{A.4a}$$

$$v_{\boldsymbol{ijkl}} = \frac{1}{V} \int_\Omega d\mathbf{r} \int_{\mathbb{R}^3} d\mathbf{r}' \frac{\phi^*_{\mathbf{k}_i,i}(\mathbf{r}) \phi_{\mathbf{k}_j,j}(\mathbf{r}) \phi^*_{\mathbf{k}_k,k}(\mathbf{r}') \phi_{\mathbf{k}_l,l}(\mathbf{r}')}{|\mathbf{r} - \mathbf{r}'|}, \tag{A.4b}$$

where Z_α is the nuclear charge of atom α, $r_{\alpha,\mathbf{r}} = |\mathbf{r} - \mathbf{r}_\alpha|$ is the distance to nucleus α at r_α, Ω is the volume of the unit cell and V is the volume of the system.

The primary object of interest in this paper is in the single-particle imaginary time Green's function $G^{\mathcal{H},\sigma}_{\boldsymbol{ij}}(\tau)$ for Hamiltonian $\mathcal{H}$ and indices $\boldsymbol{i}$ and $\boldsymbol{j}$,

$$G^{\mathcal{H},\sigma}_{\boldsymbol{ij}}(\tau) = -\frac{1}{\mathcal{Z}} Tr \left[e^{-(\beta-\tau)(\mathcal{H}-\mu N)} c_{\boldsymbol{i},\sigma} e^{-\tau(\mathcal{H}-\mu N)} c^\dagger_{\boldsymbol{j},\sigma} \right]. \tag{A.5}$$

Here $\mathcal{Z} = Tr\left[e^{-\beta(\mathcal{H}-\mu N)}\right]$ is the grand partition function, μ is the chemical potential, β is the inverse temperature and N is the number of particles in the system. We define the non-interacting Green's function as $G^{0,\sigma}_{\boldsymbol{ij}}(\tau) = G^{H^0,\sigma}_{\boldsymbol{ij}}(\tau)$, where $H^0 = \sum_{\boldsymbol{ij},\sigma} h^0_{\boldsymbol{ij}} c^\dagger_{\boldsymbol{i}\sigma} c_{\boldsymbol{j}\sigma}$, and the interacting one as $G^\sigma_{\boldsymbol{ij}}(\tau) = G^{H,\sigma}_{\boldsymbol{ij}}(\tau)$.

Translation symmetry implies that Green's functions are diagonal in reciprocal

space but dense in orbital space and can be defined as

$$G_{ij}^{\mathbf{k}\sigma}(\tau) = G_{\boldsymbol{ij}}^{\sigma}(\tau), \tag{A.6}$$

with $\mathbf{k} = \mathbf{k}_i = \mathbf{k}_j$.

The Matsubara frequency Green's function is defined through the Fourier transform

$$G_{\boldsymbol{ij}}^{\sigma}(\omega_n) = \int_0^{\beta} d\tau G_{\boldsymbol{ij}}^{\sigma}(\tau) e^{i\omega_n \tau}, \tag{A.7}$$

where $\omega_n = (2n+1)\frac{\pi}{\beta}$ is the fermionic Matsubara frequency with n integer. The self-energy is defined by the Dyson equation

$$\Sigma_{\boldsymbol{ij}}^{\sigma}(\omega_n) = \left(G_{\boldsymbol{ij}}^{0,\sigma}(\omega_n)\right)^{-1} - \left(G_{\boldsymbol{ij}}^{\sigma}(\omega_n)\right)^{-1}. \tag{A.8}$$

Knowledge of the single particle Green's function allows the computation of the spectral function or density of states as

$$G_{ij}^{\sigma}(\tau) = \int d\omega \frac{A_{ij}^{\sigma}(\omega) e^{-\tau\omega}}{1 + e^{-\beta\omega}}. \tag{A.9}$$

A.0.1 GW approximation

In a first step, we solve the system in the fully self-consistent finite temperature GW approximation introduced by Hedin. This approximation is thermodynamically consistent and conserving but neglects second-order and higher exchange terms.

The GW self-energy is given by

$$\Sigma_{ij}^{\mathbf{k},\sigma}(\omega_n) = -\frac{1}{\beta V}\sum_{\substack{m\\ \mathbf{k}',kl}}\Big[G_{lk}^{\mathbf{k}',\sigma}(\omega_n+\Omega_m)W_{ilkj}^{\mathbf{kk}'\mathbf{k}'\mathbf{k}}(\Omega_m) - \sum_{\sigma'}G_{lk}^{\mathbf{k}',\sigma'}(\omega_m)v_{ijkl}^{\mathbf{kkk}'\mathbf{k}'}\Big], \tag{A.10}$$

where $\Omega_m = \frac{2m\pi}{\beta}$ are the bosonic Matsubara frequencies and the 'screened interaction' $W_{ijkl}^{\mathbf{kk}'\mathbf{k}'}$ is defined as

$$\begin{aligned} W_{\boldsymbol{i}_1\boldsymbol{i}_2\boldsymbol{i}_3\boldsymbol{i}_4}(\Omega_n) &= v_{\boldsymbol{i}_1\boldsymbol{i}_2\boldsymbol{i}_3\boldsymbol{i}_4} + \tilde{W}_{\boldsymbol{i}_1\boldsymbol{i}_2\boldsymbol{i}_3\boldsymbol{i}_4}(\Omega_n) \\ \tilde{W}_{\boldsymbol{i}_1\boldsymbol{i}_2\boldsymbol{i}_3\boldsymbol{i}_4}(\Omega_n) &= \frac{1}{V^2} \\ &\sum_{\boldsymbol{i}_5\boldsymbol{i}_6\boldsymbol{i}_7\boldsymbol{i}_8} v_{\boldsymbol{i}_1\boldsymbol{i}_2\boldsymbol{i}_5\boldsymbol{i}_6}\Pi_{\boldsymbol{i}_5\boldsymbol{i}_6\boldsymbol{i}_7\boldsymbol{i}_8}(\Omega_n)W_{\boldsymbol{i}_7\boldsymbol{i}_8\boldsymbol{i}_3\boldsymbol{i}_4}(\Omega_n), \end{aligned} \tag{A.11}$$

with the approximate polarization operator

$$\Pi_{\boldsymbol{i}_1\boldsymbol{i}_2\boldsymbol{i}_3\boldsymbol{i}_4}(\Omega_n) = \frac{1}{\beta}\sum_m G_{\boldsymbol{i}_1\boldsymbol{i}_3}^{\sigma}(\omega_m)G_{\boldsymbol{i}_4\boldsymbol{i}_2}^{\sigma}(\omega_m+\Omega_n). \tag{A.12}$$

Eq. A.10 can be written as

$$\Sigma_{ij}^{\mathbf{k},\sigma}(\omega_n) = \Sigma_{ij}^{\mathrm{HF},\mathbf{k},\sigma} + \tilde{\Sigma}_{ij}^{\mathbf{k},\sigma}(\omega_n) \tag{A.13a}$$

$$\tilde{\Sigma}_{ij}^{\mathbf{k},\sigma}(\omega_n) = -\frac{1}{\beta V}\sum_{\substack{m\\ \mathbf{k}',kl}} G_{l,k}^{\mathbf{k}',\sigma}(\omega_n+\Omega_m)\tilde{W}_{ilkj}^{\mathbf{kk}'\mathbf{k}'\mathbf{k}}(\Omega_m), \tag{A.13b}$$

where $\Sigma_{ij}^{\mathrm{HF},\mathbf{k},\sigma}$ is the Hartree-Fock self-energy. The self-consistent GW correction to the Hartree-Fock self-energy, $\tilde{\Sigma}_{ij}^{\mathbf{k},\sigma}(\omega_n)$, contains an infinite series of 'bubble' diagrams as shown in Fig. A.1.

In our GW implementation, we use a Coulomb integral decomposition since due to its size, it is not practical to store the full four-index Coulomb integral. Several ways to employ its symmetry to decompose it are known, such as Cholesky

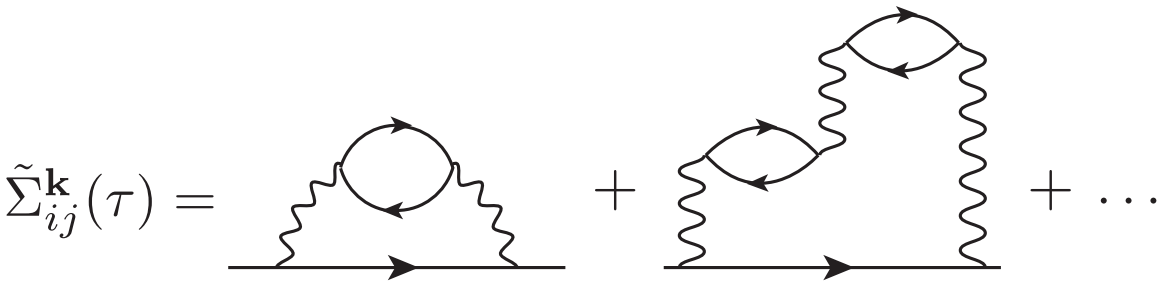


Figure A.1: Diagrams beyond the Hartree diagram in the self-consistent GW approximation. Wiggly lines denote bare interactions v, lines with arrow dressed Green's functions G.

decomposition or the resolution of identity (also known as density fitting). Here, we write $v_{\boldsymbol{i}_1\boldsymbol{i}_2\boldsymbol{i}_3\boldsymbol{i}_4} = V^Q_{\boldsymbol{i}_1\boldsymbol{i}_2} V^Q_{\boldsymbol{i}_3\boldsymbol{i}_4}$ where Q is an auxiliary index and $V^Q_{\boldsymbol{i}_1\boldsymbol{i}_2}$ is a three-point integral defined as

$$V^Q_{\boldsymbol{i}_1\boldsymbol{i}_2} = \sum_P \int_\Omega d\mathbf{r} d\mathbf{r}' \frac{\phi^*_{\boldsymbol{i}_1}(\mathbf{r})\phi_{\boldsymbol{i}_2}(\mathbf{r})\chi^{\mathbf{q}}_P(\mathbf{r}')}{|\mathbf{r}-\mathbf{r}'|} \mathbf{J}^{-\frac{1}{2}\,\mathbf{q}}_{PQ}, \tag{A.14}$$

with momentum transfer $\mathbf{q} = \mathbf{k}_1 - \mathbf{k}_2 = \mathbf{k}_3 - \mathbf{k}_4$, $\chi^{\mathbf{q}}_P(\mathbf{r}')$ an auxiliary basis function and $\mathbf{J}^{-1} = \mathbf{J}^{-\frac{1}{2}}\mathbf{J}^{-\frac{1}{2}}$ the inverse of

$$J^{\mathbf{q}}_{PQ} = \int_\Omega d\mathbf{r} d\mathbf{r}' \frac{\chi^{\mathbf{q}}_P(\mathbf{r})\chi^{\mathbf{q}}_Q(\mathbf{R}')}{|\mathbf{r}-\mathbf{r}'|}. \tag{A.15}$$

This allows to simplify Eq. A.11 to

$$\tilde{W}_{\boldsymbol{i}_1\boldsymbol{i}_2\boldsymbol{i}_3\boldsymbol{i}_4}(\Omega_n) = -\frac{1}{V}\sum_{Q,Q'} V^Q_{\boldsymbol{i}_1\boldsymbol{i}_2} \tilde{P}^{\mathbf{q}}_{QQ'}(\Omega_n) V^{Q'}_{\boldsymbol{i}_3\boldsymbol{i}_4}, \tag{A.16}$$

where the renormalized polarization matrix $\tilde{P}^{\mathbf{q}}(\Omega_n)$ is

$$\tilde{P}^{\mathbf{q}}(\Omega_n) = [\mathbb{I} - \tilde{P}^{\mathbf{q}}_0(\Omega_n)]^{-1}\tilde{P}^{\mathbf{q}}_0(\Omega_n), \tag{A.17}$$

and

$$\tilde{P}^{\mathbf{q}}_{0,Q,Q'}(\Omega_n) = \frac{1}{V} \sum_{\substack{\mathbf{k},m,\sigma \\ i_1,i_2,i_3,i_4}} V^{Q,\mathbf{k},\mathbf{k}+\mathbf{q}}_{i_1 i_2} G^{\mathbf{k},\sigma}_{i_1,i_4}(\omega_m) G^{\mathbf{k}+\mathbf{q},\sigma}_{i_3,i_2}(\omega_m + \Omega_n) V^{Q'\mathbf{k}+\mathbf{q},\mathbf{k}}_{i_3 i_4}. \tag{A.18}$$

Eq. A.13b then simplifies to

$$\tilde{\Sigma}^{\mathbf{k},\sigma}_{i_1 i_2}(\tau) = -\frac{1}{V} \sum_{\substack{\mathbf{q},i_3,i_4 \\ Q,Q'}} V^{Q,\mathbf{k}\mathbf{k}-\mathbf{q}}_{i_1,i_4} G^{\mathbf{k}-\mathbf{q},\sigma}_{i_3,i_4}(\tau) \tilde{P}^{\mathbf{q}}_{Q,Q'}(\tau) V^{Q',\mathbf{k}-\mathbf{q},\mathbf{k}}_{i_3 i_2}. \tag{A.19}$$

We diagrammatically represent this decomposition in Fig. A.2.

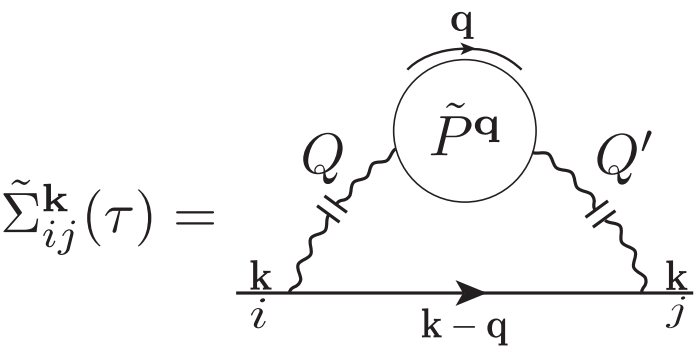


Figure A.2: Diagrams of Fig. A.1 expressed with the decomposition of Eq. A.13b. Interrupted wiggly lines denote the auxiliary basis decomposition indices Q and Q'.

A.0.2 Self-energy embedding method

GW is an approximate method with well known limitations. To capture correlation effects beyond the GW approximation, either high-order diagrammatic methods or quantum embedding methods can be used. Embedding theories that are Φ-derivable and based on diagrammatic expansions such as DMFT, GW+EDMFT, SEET, or self-energy functional theory aim to systematically improve low-order perturbative

results. These embedding theories satisfy conservation laws and are thermodynamically consistent.

Here, we briefly summarize the SEET equations used by in this paper. The real space Green's function and the lattice (k-space) Green's function are related by the Fourier transform

$$G_{ij}^{\mathbf{R}\mathbf{R}'}(\omega_n) = \frac{1}{V}\sum_{k} e^{i\mathbf{k}\mathbf{R}} G_{ij}^{\mathbf{k}}(\omega_n) e^{-i\mathbf{k}\mathbf{R}'}. \tag{A.20}$$

The GW momentum resolved Green's function of the entire lattice is defined as

$$(G^{\mathrm{GW}}(\omega_n))^{\mathbf{k}} = \left[(\omega_n + \mu)\mathbb{I} - h^{0,\mathbf{k}} - (\Sigma^{\mathrm{GW}})^{\mathbf{k}}\right]^{-1}, \tag{A.21}$$

where $(\Sigma^{\mathrm{GW}})^{\mathbf{k}} = (\Sigma_{\infty}^{\mathrm{GW}})^{\mathbf{k}} + (\Sigma^{\mathrm{GW}}(\omega))^{\mathbf{k}}$. As a result of embedding procedure, we define a lattice Green's function in the following way

$$(G(\omega_n))^{\mathbf{k}} = \left[(\omega + \mu)\mathbb{I} - h^{0,\mathbf{k}} - \Sigma^{\mathbf{k}}\right]^{-1}, \tag{A.22}$$

where

$$\Sigma_{ij}^{\mathbf{k}} = (\Sigma^{GW})_{ij}^{\mathbf{k}} + \sum_{A} \left((\Sigma_A^{\mathrm{imp}})_{ij} - (\Sigma_A^{\text{DC-GW}})_{ij} \right) \delta_{(ij)\in A} \tag{A.23}$$

with $\Sigma^{\mathrm{imp}} = \Sigma_{\infty}^{\mathrm{imp}} + \Sigma^{\mathrm{imp}}(\omega_n)$ containing non-perturbatively added self-energy diagrams and $\Sigma^{\text{DC-GW}} = \Sigma_{\infty}^{\text{DC-GW}} + \Sigma^{\text{DC-GW}}(\omega_n)$ subtracting those diagrams that are contained both in the GW solution and the non-perturbative construction. Subsets A of impurity orbitals with indices $ij \in A$, sometimes also called active orbitals, are defined as groups of the most physically relevant orbitals for the problem that have correlations that are necessary to be included at a higher than perturbative level.

To define the self-consistency condition used in SEET we perform Fourier transform

of $(G(\omega))^{\mathbf{k}}$, $\Sigma^{\mathbf{k}}$, and $h^{0,\mathbf{k}}$ from momentum to real space obtaining $G^{\mathbf{RR'}}$, $\Sigma^{\mathbf{RR'}}$, and $h^{0,\mathbf{RR'}}$. The Fourier transform results in the following structure of the self-energy matrix in the real space

$$\begin{aligned}\Sigma_{ij}^{\mathbf{RR'}} &= (\Sigma^{GW})_{ij}^{\mathbf{RR'}} \\ &+ \sum_{A} \left((\Sigma_A^{\mathrm{imp}})_{ij} - (\Sigma_A^{\mathrm{DC}})_{ij} \right) \delta_{\mathbf{RR'}} \delta_{(ij)\in A},\end{aligned} \tag{A.24}$$

for unit cells away from central cell ($\mathbf{R} \neq \mathbf{R'}$) the self-energies are treated at the weakly correlated level $\Sigma_{ij}^{\mathbf{RR'}} = (\Sigma^{GW})_{ij}^{\mathbf{RR'}}$ while the local, central cell self-energy for $\mathbf{R} = \mathbf{R'}$ includes non-perturbative corrections $(\Sigma_A^{\mathrm{imp}})_{ij}$ for every orbital group A. This leads us to a definition of an embedding condition in SEET, where we apply the block-matrix inversions of real space quantities and absorb all terms containing contributions connecting orbitals in A to the remainder of the system in the matrix $\Delta_{ij}^{A}(\omega)$ in the following way

$$(G(\omega_n))_{ij\in A}^{\mathbf{RR}} = \left[(\omega_n + \mu)\mathbb{I} - h_{ij\in A}^{0,\mathbf{RR}} - \Sigma_{ij\in A}^{\mathbf{RR}} - \Delta_{ij}^{A}(\omega_n) \right]^{-1}. \tag{A.25}$$

The hybridization matrix $\Delta_{ij}^{A}(\omega_n)$ arises since an inverse of a subset is not equal to a subset of an inverse, namely
$(G(\omega_n))_{ij\in A}^{\mathbf{RR}} \neq [(G(\omega_n))^{\mathbf{RR'}})^{-1}]_{ij\in A}^{\mathbf{RR}} = \left[(\omega_n + \mu)\mathbb{I} - h_{ij\in A}^{0,\mathbf{RR}} - \Sigma_{ij\in A}^{\mathbf{RR}} \right]^{-1}$. Note that Eq. A.25 can further be rewritten as

$$\begin{aligned}[(G(\omega_n))_{ij\in A}^{\mathbf{RR}}]^{-1} &= (i\omega_n + \mu)\mathbb{I} - \tilde{h}_{ij\in A}^{0,\mathbf{RR}} + \\ &- \Sigma_{ij\in A}^{\mathrm{corr},\mathbf{RR}}(\omega_n) - \Sigma_{ij\in A}^{\mathrm{imp}} - \Delta_{ij}^{A}(\omega_n),\end{aligned} \tag{A.26}$$

where $\tilde{h}_{ij}^{0,\mathbf{RR}} = h_{ij}^{0,\mathbf{RR}} + (\Sigma^{\mathrm{GW}})_{\infty,ij}^{\mathbf{RR}} - \Sigma_{\infty,ij}^{\mathrm{DC}}$ is the renormalized noninteracting Hamiltonian, and $\Sigma_{ij}^{\mathrm{corr},\mathbf{RR}}(\omega_n) = (\Sigma^{\mathrm{GW}})_{ij}^{\mathbf{RR}}(\omega_n) - \Sigma_{ij}^{\mathrm{DC}}(\omega_n)$ is the local correction from the weakly correlated method.

We emphasize that, in SEET, the substantial contribution of $\Sigma^{\mathrm{corr},\mathbf{RR}}_{ij\in A}(\omega_n)$ to the local correlated orbitals is included explicitly in the real space self-consistency condition in Eq. A.25 and is not included as a part of hybridization as done in the GW+DMFT schemes. These contributions stem from GW diagrams that have both external legs i and j in the active space but contain one or more internal indices on the remaining orbitals. Furthermore, the explicit treatment of $\Sigma^{\mathrm{corr},\mathbf{RR}}_{ij\in A}(\omega_n)$ prevents us from observing non-causality problems with hybridization. $\Delta^A_{ij}(\omega_n)$ as defined in Eq. A.25 is always causal.

To evaluate $\Sigma^{\mathrm{imp}}_{ij\in A}$, we define the auxiliary propagator

$$\mathcal{G}^{-1}_A(\omega_n) = \mathcal{G}^{0,-1}_A(\omega_n) - \Sigma^{\mathrm{imp}}_{ij\in A}, \tag{A.27}$$

where the zeroth order $\mathcal{G}^{0,-1}_A(\omega_n)$ is defined as

$$\mathcal{G}^{0,-1}_A(\omega_n) = (i\omega_n + \mu)\delta_{ij} - \tilde{h}^{0,\mathbf{RR}}_{ij\in A} - \Delta^A_{ij}(\omega_n). \tag{A.28}$$

As realized in the context of DMFT, a propagator of the form of Eq. A.27 can be obtained by solving the quantum impurity model with impurity orbitals defined as the active orbitals from a space A. In SEET, the two-body interactions in the impurity remain the bare, unchanged interactions of the original lattice Hamiltonian, since screening is included by the explicit treatment of $\Sigma^{\mathrm{corr},\mathbf{RR}}_{ij\in A}(\omega_n)$ at the level of the embedding condition and Eq. A.26.

The fact that the bare interactions do not need to be adjusted in the impurity model is a major difference to formulations of GW+EDMFT. The GW+EDMFT double counting correction due to the presence of screened $W^{\mathrm{imp}}(\omega_n)$ removes local correction to the self-energy from the weakly correlated method, therefore $\Sigma^{corr,\mathbf{RR}}_{ij}(\omega_n) \equiv 0$.This GW+EDMFT construction containing $W^{\mathrm{imp}}(\omega_n)$ leads to an impurity model with a different hybridization and noninteracting Hamiltonian and,

as the model needs to take into account correlations outside the active space accordingly, to a rescaling of the interactions. However, while operationally different, both GW+EDMFT and SEET are consistent, conserving, and contain RPA screening by GW diagrams.

In practice, our method starts from a self-consistent finite temperature GW solution of the lattice problem. It then proceeds by solving all independent impurity problems for the different disjoint subspaces A independently. The non-perturbative solution of $\Sigma_{ij}^{\mathrm{imp}}$ is used to update the lattice self-energy and the Green's function from Eq. A.23 and A.22, followed by a new calculation of the real space Green's function and hybridization (Eq. A.25) and a subsequent solution of the impurity model. In principle, after obtaining the self-consistent solution of Eq. A.25, the GW solution would need to be iterated again. This has not been done in this work.

BIBLIOGRAPHY

BIBLIOGRAPHY

Abrikosov, A. A., I. Dzyaloshinskii, L. P. Gorkov, and R. A. Silverman (1975), *Methods of quantum field theory in statistical physics*, Dover, New York, NY.

Adams, B. M., et al. (2014), Dakota, a multilevel parallel object-oriented framework for design optimization, parameter estimation, uncertainty quantification, and sensitivity analysis: Version 6.11 user's manual.

Almlof, J., and P. R. Taylor (1987), General contraction of gaussian basis sets. i. atomic natural orbitals for first- and second-row atoms, *The Journal of Chemical Physics*, *86*(7), 4070–4077, doi:10.1063/1.451917.

Andrae, D., U. Häußermann, M. Dolg, H. Stoll, and H. Preuß (1990), Energy-adjustedab initio pseudopotentials for the second and third row transition elements, *Theoretica chimica acta*, *77*(2), 123–141, doi:10.1007/BF01114537.

Aryasetiawan, F., and O. Gunnarsson (1998a), The gw method, *Reports on Progress in Physics*, *61*(3), 237.

Aryasetiawan, F., and O. Gunnarsson (1998b), ThebiGW/i/bmethod, *Reports on Progress in Physics*, *61*(3), 237–312, doi:10.1088/0034-4885/61/3/002.

Aryasetiawan, F., and O. Gunnarsson (1998c), The gw method, *Reports on Progress in Physics*, *61*(3), 237.

Aulbur, W. G., L. Jönsson, and J. W. Wilkins (2000), Quasiparticle calculations in solids, *Journal of Physics C: Solid State Physics*, *54*, 1–218.

Ayala, P. Y., K. N. Kudin, and G. E. Scuseria (2001), Atomic orbital laplace-transformed second-order møller–plesset theory for periodic systems, *The Journal of Chemical Physics*, *115*(21), 9698–9707, doi:10.1063/1.1414369.

Becke, A. D. (1988), Density-functional exchange-energy approximation with correct asymptotic behavior, *Phys. Rev. A*, *38*, 3098–3100, doi:10.1103/PhysRevA.38.3098.

Becke, A. D. (1993), Density-functional thermochemistry. iii. the role of exact exchange, *The Journal of Chemical Physics*, *98*(7), 5648–5652, doi:10.1063/1.464913.

Biermann, S., F. Aryasetiawan, and A. Georges (2003), First-principles approach to the electronic structure of strongly correlated systems: Combining the *gw* approximation and dynamical mean-field theory, *Phys. Rev. Lett.*, *90*, 086,402, doi:10.1103/PhysRevLett.90.086402.

Blaha, P., K. Schwarz, F. Tran, R. Laskowski, G. K. H. Madsen, and L. D. Marks (2020a), Wien2k: An apw+lo program for calculating the properties of solids, *The Journal of Chemical Physics*, *152*(7), 074,101, doi:10.1063/1.5143061.

Blaha, P., K. Schwarz, F. Tran, R. Laskowski, G. K. H. Madsen, and L. D. Marks (2020b), Wien2k: An apw+lo program for calculating the properties of solids, *The Journal of Chemical Physics*, *152*(7), 074,101, doi:10.1063/1.5143061.

Boehnke, L., F. Nilsson, F. Aryasetiawan, and P. Werner (2016), When strong correlations become weak: Consistent merging of *gw* and dmft, *Phys. Rev. B*, *94*, 201,106, doi:10.1103/PhysRevB.94.201106.

Boys, S. F. (1950), Electronic wave functions. i. a general method of calculation for the stationary states of any molecular system, *Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences*, *200*(1063), 542–554.

Cao, X., and M. Dolg (2011), Pseudopotentials and modelpotentials, *WIREs Computational Molecular Science*, *1*(2), 200–210, doi:https://doi.org/10.1002/wcms.28.

Caruso, F., P. Rinke, X. Ren, A. Rubio, and M. Scheffler (2013), Self-consistent *gw*: All-electron implementation with localized basis functions, *Phys. Rev. B*, *88*, 075,105, doi:10.1103/PhysRevB.88.075105.

Caruso, F., M. Dauth, M. J. van Setten, and P. Rinke (2016), Benchmark of gw approaches for the gw100 test set, *Journal of Chemical Theory and Computation*, *12*(10), 5076–5087, doi:10.1021/acs.jctc.6b00774.

Cheney, E. W., and D. R. Kincaid (2007), *Numerical Mathematics and Computing*, 6th ed., Brooks/Cole Publishing Co., USA.

Cox, S., and J. Fry (1977), Convergence of reciprocal-lattice sums: Exact formulas for the ewald method, *Journal of Computational Physics*, *23*(1), 42–52, doi:https://doi.org/10.1016/0021-9991(77)90086-9.

CRYSTAL (2021), Crystal - basis sets library, https://www.crystal.unito.it/basis-sets.php, accessed: 2020-11-17.

Csonka, G. I., J. P. Perdew, A. Ruzsinszky, P. H. T. Philipsen, S. Lebègue, J. Paier, O. A. Vydrov, and J. G. Ángyán (2009), Assessing the performance of recent density functionals for bulk solids, *Phys. Rev. B*, *79*, 155,107, doi:10.1103/PhysRevB.79.155107.

Daga, L. E., B. Civalleri, and L. Maschio (2020), Gaussian basis sets for crystalline solids: All-purpose basis set libraries vs system-specific optimizations, *Journal of Chemical Theory and Computation*, *16*(4), 2192–2201, doi:10.1021/acs.jctc.9b01004, pMID: 32212698.

Dahlen, N. E., and R. van Leeuwen (2005), Self-consistent solution of the dyson equation for atoms and molecules within a conserving approximation, *The Journal of Chemical Physics*, *122*(16), 164102, doi:http://dx.doi.org/10.1063/1.1884965.

Davidson, E. R., and D. Feller (1986), Basis set selection for molecular calculations, *Chemical Reviews*, *86*(4), 681–696, doi:10.1021/cr00074a002.

Del Ben, M., J. Hutter, and J. VandeVondele (2012), Second-order møller–plesset perturbation theory in the condensed phase: An efficient and massively parallel gaussian and plane waves approach, *Journal of Chemical Theory and Computation*, *8*(11), 4177–4188, doi:10.1021/ct300531w, pMID: 26605583.

Delaney, K., P. García-González, A. Rubio, P. Rinke, and R. W. Godby (2004), Comment on "band-gap problem in semiconductors revisited: Effects of core states and many-body self-consistency", *Phys. Rev. Lett.*, *93*, 249,701, doi:10.1103/PhysRevLett.93.249701.

Dolg, M., and X. Cao (2012), Relativistic pseudopotentials: Their development and scope of applications, *Chemical Reviews*, *112*(1), 403–480, doi:doi:10.1021/cr2001383.

Dovesi, R., M. Causa', R. Orlando, C. Roetti, and V. R. Saunders (1990), Ab initio approach to molecular crystals: A periodic hartree–fock study of crystalline urea, *The Journal of Chemical Physics*, *92*(12), 7402–7411, doi:10.1063/1.458592.

Dovesi, R., et al. (2017), *CRYSTAL17 User's Manual*, University of Torino, Torino.

Dovesi, R., et al. (2018), Quantum-mechanical condensed matter simulations with crystal, *WIREs Computational Molecular Science*, *8*(4), e1360, doi:https://doi.org/10.1002/wcms.1360.

Ehrenreich, H., F. Seitz, and D. Turnbull (1978), *Solid State Physics*, ISSN, Elsevier Science.

Gatti, C., V. R. Saunders, and C. Roetti (1994), Crystal field effects on the topological properties of the electron density in molecular crystals: The case of urea, *The Journal of Chemical Physics*, *101*(12), 10,686–10,696, doi:10.1063/1.467882.

Georges, A., and G. Kotliar (1992), Hubbard model in infinite dimensions, *Phys. Rev. B*, *45*, 6479–6483, doi:10.1103/PhysRevB.45.6479.

Georges, A., G. Kotliar, W. Krauth, and M. J. Rozenberg (1996), Dynamical mean-field theory of strongly correlated fermion systems and the limit of infinite dimensions, *Rev. Mod. Phys.*, *68*, 13–125, doi:10.1103/RevModPhys.68.13.

Godby, R. W., M. Schlüter, and L. J. Sham (1988), Self-energy operators and exchange-correlation potentials in semiconductors, *Phys. Rev. B*, *37*, 10,159–10,175, doi:10.1103/PhysRevB.37.10159.

Golub, G. H., and C. F. Van Loan (1996), *Matrix Computations (3rd Ed.)*, Johns Hopkins University Press, USA.

Gull, E., S. Iskakov, I. Krivenko, A. A. Rusakov, and D. Zgid (2018), Chebyshev polynomial representation of imaginary-time response functions, *Phys. Rev. B*, *98*, 075,127, doi:10.1103/PhysRevB.98.075127.

Hedin, L. (1965), New method for calculating the one-particle green's function with application to the electron-gas problem, *Phys. Rev.*, *139*, A796–A823, doi:10.1103/PhysRev.139.A796.

Hedin, L. (1999), On correlation effects in electron spectroscopies and the gw approximation, *Journal of Physics: Condensed Matter*, *1999*(11), 489–528.

Helgaker, T., P. Jorgensen, and J. Olsen (2000a), *Atomic Basis Functions*, chap. 6, pp. 201–255, John Wiley & Sons, Ltd, doi:https://doi.org/10.1002/9781119019572.ch6.

Helgaker, T., P. Jorgensen, and J. Olsen (2000b), *Molecular Electronic-Structure Theory*, John Wiley & Sons, doi:10.1002/9781119019572.

Hestenes, M. R., and E. Stiefel (1952), Methods of conjugate gradients for linear systems, *Journal of Research of the National Bureau of Standards*, *49*(6), 409, doi:10.6028/jres.049.044.

Heyd, J., J. E. Peralta, G. E. Scuseria, and R. L. Martin (2005), Energy band gaps and lattice parameters evaluated with the heyd-scuseria-ernzerhof screened hybrid functional, *The Journal of Chemical Physics*, *123*(17), 174,101, doi:10.1063/1.2085170.

Hohenberg, P., and W. Kohn (1964a), Inhomogeneous electron gas, *Phys. Rev.*, *136*, B864–B871, doi:10.1103/PhysRev.136.B864.

Hohenberg, P., and W. Kohn (1964b), Inhomogeneous electron gas, *Phys. Rev.*, *136*, B864–B871, doi:10.1103/PhysRev.136.B864.

Holm, B., and U. von Barth (1998), Fully self-consistent GW self-energy of the electron gas, *Phys. Rev. B*, *57*, 2108–2117, doi:10.1103/PhysRevB.57.2108.

Hooke, R., and T. A. Jeeves (1961), Direct search solution of numerical and statistical problems, *Journal of the ACM*, *8*(2), 212–229, doi:10.1145/321062.321069.

Huzinaga, S. (1985), Basis sets for molecular calculations, *Computer Physics Reports*, *2*(6), 281–339, doi:https://doi.org/10.1016/0167-7977(85)90003-6.

Ihm, J., A. Zunger, and M. L. Cohen (1979), Momentum-space formalism for the total energy of solids, *Journal of Physics C: Solid State Physics*, *12*(21), 4409–4422, doi:10.1088/0022-3719/12/21/009.

Imada, M., A. Fujimori, and Y. Tokura (1998), Metal-insulator transitions, *Rev. Mod. Phys.*, *70*, 1039–1263, doi:10.1103/RevModPhys.70.1039.

Iskakov, S., C.-N. Yeh, E. Gull, and D. Zgid (2020a), Ab initio self-energy embedding for the photoemission spectra of nio and mno, *Phys. Rev. B*, *102*, 085,105, doi:10.1103/PhysRevB.102.085105.

Iskakov, S., C.-N. Yeh, E. Gull, and D. Zgid (2020b), Ab initio self-energy embedding for the photoemission spectra of nio and mno, *Phys. Rev. B*, *102*, 085,105, doi:10.1103/PhysRevB.102.085105.

Iskakov, S., C.-N. Yeh, E. Gull, and D. Zgid (2020c), iab initio/i self-energy embedding for the photoemission spectra of NiO and MnO, *Physical Review B*, *102*(8), doi:10.1103/physrevb.102.085105.

Jensen, F. (2006), *Introduction to Computational Chemistry*, John Wiley & Sons, Inc., Hoboken, NJ, USA.

Kananenka, A. A., and D. Zgid (2017), Combining density functional theory and green's function theory: Range-separated, nonlocal, dynamic, and orbital-dependent hybrid functional, *J. Chem. Theory Comput.*, *13*(0), 5317, doi:10.1021/acs.jctc.7b00701.

Kananenka, A. A., E. Gull, and D. Zgid (2015), Systematically improvable multiscale solver for correlated electron systems, *Phys. Rev. B*, *91*, 121,111(R), doi:10.1103/PhysRevB.91.121111.

Kananenka, A. A., J. J. Phillips, and D. Zgid (2016a), Efficient temperature-dependent green's functions methods for realistic systems: Compact grids for orthogonal polynomial transforms, *J. Chem. Theory Comput.*, *12*(2), 564, doi:10.1021/acs.jctc.5b00884.

Kananenka, A. A., A. R. Welden, T. N. Lan, E. Gull, and D. Zgid (2016b), Efficient temperature-dependent green's function methods for realistic systems: Using cubic spline interpolation to approximate matsubara green's functions, *J. Chem. Theory Comput.*, *12*(5), 2250–2259, doi:10.1021/acs.jctc.6b00178.

Karpfen, A. (1979), Near hartree-fock limit study of the linear chain of hydrogen atoms, *Chemical Physics Letters*, *61*(2), 363–366, doi:https://doi.org/10.1016/0009-2614(79)80663-6.

Klein, R. A., et al. (2019), High-pressure synthesis of the $BiVo_3$ perovskite, *Phys. Rev. Materials*, *3*, 064,411, doi:10.1103/PhysRevMaterials.3.064411.

Knizia, G., and G. K.-L. Chan (2012), Density matrix embedding: A simple alternative to dynamical mean-field theory, *Phys. Rev. Lett.*, *109*, 186,404, doi:10.1103/PhysRevLett.109.186404.

Knizia, G., and G. K.-L. Chan (2013), Density matrix embedding: A strong-coupling quantum embedding theory, *Journal of Chemical Theory and Computation*, *9*(3), 1428–1432, doi:10.1021/ct301044e, pMID: 26587604.

Kohn, W., and L. J. Sham (1965), Self-consistent equations including exchange and correlation effects, *Phys. Rev.*, *140*, A1133–A1138, doi:10.1103/PhysRev.140.A1133.

Kotliar, G., S. Y. Savrasov, K. Haule, V. S. Oudovenko, O. Parcollet, and C. A. Marianetti (2006), Electronic structure calculations with dynamical mean-field theory, *Rev. Mod. Phys.*, *78*, 865–951, doi:10.1103/RevModPhys.78.865.

Koval, P., D. Foerster, and D. Sánchez-Portal (2014), Fully self-consistent *gw* and quasiparticle self-consistent *gw* for molecules, *Phys. Rev. B*, *89*, 155,417, doi:10.1103/PhysRevB.89.155417.

Kresse, G., and J. Furthmüller (1996a), Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, *Phys. Rev. B*, *54*, 11,169–11,186, doi:10.1103/PhysRevB.54.11169.

Kresse, G., and J. Furthmüller (1996b), Efficient iterative schemes for ab initio total-energy calculations using a plane-wave basis set, *Phys. Rev. B*, *54*, 11,169–11,186, doi:10.1103/PhysRevB.54.11169.

Ku, W., and A. G. Eguiluz (2002), Band-gap problem in semiconductors revisited: Effects of core states and many-body self-consistency, *Phys. Rev. Lett.*, *89*, 126,401, doi:10.1103/PhysRevLett.89.126401.

Kudin, K. N., and G. E. Scuseria (2000), Linear-scaling density-functional theory with gaussian orbitals and periodic boundary conditions: Efficient evaluation of energy and forces via the fast multipole method, *Phys. Rev. B*, *61*, 16,440–16,453, doi:10.1103/PhysRevB.61.16440.

Kuhne, T. D., et al. (2020), Cp2k: An electronic structure and molecular dynamics software package - quickstep: Efficient and accurate electronic structure calculations, *The Journal of Chemical Physics*, *152*(19), 194,103, doi:10.1063/5.0007045.

Lan, T. N., and D. Zgid (2017), Generalized self-energy embedding theory, *The Journal of Physical Chemistry Letters*, *8*(10), 2200–2205, doi:10.1021/acs.jpclett.7b00689.

Lan, T. N., A. A. Kananenka, and D. Zgid (2016), Rigorous ab initio quantum embedding for quantum chemistry using green's function theory: Screened interaction, nonlocal self-energy relaxation, orbital basis, and chemical accuracy, *Journal of Chemical Theory and Computation*, *12*(10), 4856–4870, doi:10.1021/acs.jctc.6b00638.

Lan, T. N., A. Shee, J. Li, E. Gull, and D. Zgid (2017a), Testing self-energy embedding theory in combination with GW, *Physical Review B*, *96*(15), doi:10.1103/physrevb.96.155106.

Lan, T. N., A. Shee, J. Li, E. Gull, and D. Zgid (2017b), Testing self-energy embedding theory in combination with gw, *Phys. Rev. B*, *96*, 155,106, doi:10.1103/PhysRevB.96.155106.

Laun, J., D. Vilela Oliveira, and T. Bredow (2018), Consistent gaussian basis sets of double- and triple-zeta valence with polarization quality of the fifth period for solid-state calculations, *Journal of Computational Chemistry*, *39*(19), 1285–1290, doi:https://doi.org/10.1002/jcc.25195.

Lee, C., W. Yang, and R. G. Parr (1988), Development of the colle-salvetti correlation-energy formula into a functional of the electron density, *Phys. Rev. B*, *37*, 785–789, doi:10.1103/PhysRevB.37.785.

Lehmann, H. (1954), On the Properties of propagation functions and renormalization contants of quantized fields, *Nuovo Cim.*, *11*, 342–357, doi:10.1007/BF02783624.

Lu, J.-B., D. C. Cantu, M.-T. Nguyen, J. Li, V.-A. Glezakou, and R. Rousseau (2019), Norm-conserving pseudopotentials and basis sets to explore lanthanide chemistry in complex environments, *Journal of Chemical Theory and Computation*, *15*(11), 5987–5997, doi:10.1021/acs.jctc.9b00553, pMID: 31580663.

Luo, W., S. Ismail-Beigi, M. L. Cohen, and S. G. Louie (2002), Quasiparticle band structure of zns and znse, *Phys. Rev. B*, *66*, 195,215, doi:10.1103/PhysRevB.66.195215.

Luttinger, J. M., and J. C. Ward (1960), Ground-state energy of a many-fermion system. ii, *Phys. Rev.*, *118*, 1417–1427, doi:10.1103/PhysRev.118.1417.

March, N. (1999), *Electron Correlations In The Solid State*, World Scientific Publishing Company.

Martin, R. M. (2004), *Determination of electronic structure: the three basic methods*, p. 233–312, Cambridge University Press, doi:10.1017/CBO9780511805769.014.

Marzari, N., A. A. Mostofi, J. R. Yates, I. Souza, and D. Vanderbilt (2012), Maximally localized wannier functions: Theory and applications, *Rev. Mod. Phys.*, *84*, 1419–1475, doi:10.1103/RevModPhys.84.1419.

Maschio, L., D. Usvyat, F. R. Manby, S. Casassa, C. Pisani, and M. Schütz (2007), Fast local-mp2 method with density-fitting for crystals. i. theory and algorithms, *Phys. Rev. B*, *76*, 075,101, doi:10.1103/PhysRevB.76.075101.

Maschio, L., D. Usvyat, M. Schütz, and B. Civalleri (2010), Periodic local møller–plesset second order perturbation theory method applied to molecular crystals: Study of solid nh3 and co2 using extended basis sets, *The Journal of Chemical Physics*, *132*(13), 134,706, doi:10.1063/1.3372800.

Matsubara, T. (1955), A New Approach to Quantum-Statistical Mechanics, *Progress of Theoretical Physics*, *14*(4), 351–378, doi:10.1143/PTP.14.351.

Morales, M. A., and F. D. Malone (2020), Accelerating the convergence of auxiliary-field quantum monte carlo in solids with optimized gaussian basis sets, *The Journal of Chemical Physics*, *153*(19), 194,111, doi:10.1063/5.0025390.

Mortensen, J. J., L. B. Hansen, and K. W. Jacobsen (2005), Real-space grid implementation of the projector augmented wave method, *Phys. Rev. B*, *71*, 035,109, doi:10.1103/PhysRevB.71.035109.

Nagy, B., and F. Jensen (2017), *Basis Sets in Quantum Chemistry*, chap. 3, pp. 93–149, John Wiley & Sons, Ltd, doi:doi.org/10.1002/9781119356059.ch3.

Nilsson, F., L. Boehnke, P. Werner, and F. Aryasetiawan (2017), Multitier self-consistent gw + EDMFT, *Phys. Rev. Materials*, *1*, 043,803, doi:10.1103/PhysRevMaterials.1.043803.

Parrish, R. M., et al. (2017), Psi4 1.1: An open-source electronic structure program emphasizing automation, advanced libraries, and interoperability, *Journal of Chemical Theory and Computation*, *13*(7), 3185–3197, doi:10.1021/acs.jctc.7b00174, pMID: 28489372.

Peintinger, M. F., D. V. Oliveira, and T. Bredow (2013), Consistent gaussian basis sets of triple-zeta valence with polarization quality for solid-state calculations, *Journal of Computational Chemistry*, *34*(6), 451–459, doi:10.1002/jcc.23153.

Perdew, J. P., J. A. Chevary, S. H. Vosko, K. A. Jackson, M. R. Pederson, D. J. Singh, and C. Fiolhais (1992), Atoms, molecules, solids, and surfaces: Applications of the generalized gradient approximation for exchange and correlation, *Phys. Rev. B*, *46*, 6671–6687, doi:10.1103/PhysRevB.46.6671.

Perdew, J. P., K. Burke, and M. Ernzerhof (1997), Generalized gradient approximation made simple [phys. rev. lett. 77, 3865 (1996)], *Phys. Rev. Lett.*, *78*, 1396–1396, doi:10.1103/PhysRevLett.78.1396.

Perdew, J. P., A. Ruzsinszky, G. I. Csonka, O. A. Vydrov, G. E. Scuseria, L. A. Constantin, X. Zhou, and K. Burke (2009), Erratum: Restoring the density-gradient expansion for exchange in solids and surfaces [phys. rev. lett.

100, 136406 (2008)], *Phys. Rev. Lett.*, *102*, 039,902, doi:10.1103/PhysRevLett.102.039902.

Petersson, G. A., S. Zhong, J. A. Montgomery, and M. J. Frisch (2003), On the optimization of Gaussian basis sets, *J. Chem. Phys.*, *118*(3), 1101–1109, doi:10.1063/1.1516801.

Petocchi, F., F. Nilsson, F. Aryasetiawan, and P. Werner (2020), Screening from e_g states and antiferromagnetic correlations in $d^{(1,2,3)}$ perovskites: A gw + EDMFT investigation, *Phys. Rev. Research*, *2*, 013,191, doi:10.1103/PhysRevResearch.2.013191.

Phillips, J. J., and D. Zgid (2014), Communication: The description of strong correlation within self-consistent green's function second-order perturbation theory, *J. Chem. Phys.*, *140*(24), 241101, doi:http://dx.doi.org/10.1063/1.4884951.

Phillips, J. J., A. A. Kananenka, and D. Zgid (2015), Fractional charge and spin errors in self-consistent green's function theory, *J. Chem. Phys.*, *142*(19), 194108, doi:http://dx.doi.org/10.1063/1.4921259.

Pisani, C., M. Busso, G. Capecchi, S. Casassa, R. Dovesi, L. Maschio, C. Zicovich-Wilson, and M. Schütz (2005), Local-mp2 electron correlation method for nonconducting crystals, *The Journal of Chemical Physics*, *122*(9), 094,113, doi:10.1063/1.1857479.

Pritchard, B. P., D. Altarawy, B. Didier, T. D. Gibson, and T. L. Windus (2019), New basis set exchange: An open, up-to-date resource for the molecular sciences community, *Journal of Chemical Information and Modeling*, *59*(11), 4814–4820, doi:10.1021/acs.jcim.9b00725.

Rusakov, A. A., and D. Zgid (2016), Self-consistent second-order green's function perturbation theory for periodic systems, *J. Chem. Phys.*, *144*(5), 054106, doi:http://dx.doi.org/10.1063/1.4940900.

Rusakov, A. A., J. J. Phillips, and D. Zgid (2014), Local hamiltonians for quantitative green's function embedding methods, *The Journal of Chemical Physics*, *141*(19), 194105, doi:http://dx.doi.org/10.1063/1.4901432.

Rusakov, A. A., S. Iskakov, L. N. Tran, and D. Zgid (2018), Self-energy embedding theory (SEET) for periodic systems, *Journal of Chemical Theory and Computation*, *15*(1), 229–240, doi:10.1021/acs.jctc.8b00927.

Rusakov, A. A., S. Iskakov, L. N. Tran, and D. Zgid (2019), Self-energy embedding theory (seet) for periodic systems, *Journal of Chemical Theory and Computation*, *15*(1), 229–240, doi:10.1021/acs.jctc.8b00927.

Schuchardt, K. L., B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, and T. L. Windus (2007), Basis set exchange: A community database for computational sciences, *Journal of Chemical Information and Modeling*, *47*(3), 1045–1052, doi:10.1021/ci600510j.

Slater, J. C. (1930), Atomic shielding constants, *Phys. Rev.*, *36*, 57–64, doi:10.1103/PhysRev.36.57.

Stan, A., N. E. Dahlen, and R. van Leeuwen (2006), Fully self-consistent gw calculations for atoms and molecules, *EPL (Europhysics Letters)*, *76*(2), 298.

Suhai, S., P. Bagus, and J. Ladik (1982), An error analysis for hartree-fock crystal orbital calculations, *Chemical Physics*, *68*(3), 467–471, doi:https://doi.org/10.1016/0301-0104(82)87055-9.

Tosoni, S., C. Tuma, J. Sauer, B. Civalleri, and P. Ugliengo (2007), A comparison between plane wave and gaussian-type orbital basis sets for hydrogen bonded systems: Formic acid as a test case, *The Journal of Chemical Physics*, *127*(15), 154,102, doi:10.1063/1.2790019.

Tran, L. N., S. Iskakov, and D. Zgid (2018), Spin-unrestricted self-energy embedding theory, *The Journal of Physical Chemistry Letters*, *9*(15), 4444–4450, doi:10.1021/acs.jpclett.8b01754.

Usvyat, D., L. Maschio, F. R. Manby, S. Casassa, M. Schütz, and C. Pisani (2007), Fast local-mp2 method with density-fitting for crystals. ii. test calculations and application to the carbon dioxide crystal, *Phys. Rev. B*, *76*, 075,102, doi:10.1103/PhysRevB.76.075102.

Usvyat, D., B. Civalleri, L. Maschio, R. Dovesi, C. Pisani, and M. Schütz (2011), Approaching the theoretical limit in periodic local mp2 calculations with atomic-orbital basis sets: The case of lih, *The Journal of Chemical Physics*, *134*(21), 214,105, doi:10.1063/1.3595514.

van Schilfgaarde, M., T. Kotani, and S. Faleev (2006), Quasiparticle self-consistent *gw* theory, *Phys. Rev. Lett.*, *96*, 226,402, doi:10.1103/PhysRevLett.96.226402.

VandeVondele, J., and J. Hutter (2007a), Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases, *The Journal of Chemical Physics*, *127*(11), 114,105, doi:10.1063/1.2770708.

VandeVondele, J., and J. Hutter (2007b), Gaussian basis sets for accurate calculations on molecular systems in gas and condensed phases, *The Journal of Chemical Physics*, *127*(11), 114,105, doi:10.1063/1.2770708.

Vilela Oliveira, D., J. Laun, M. F. Peintinger, and T. Bredow (2019), Bsse-correction scheme for consistent gaussian basis sets of double- and triple-zeta valence with polarization quality for solid-state calculations, *Journal of Computational Chemistry*, *40*(27), 2364–2376, doi:10.1002/jcc.26013.

Welborn, M., T. Tsuchimochi, and T. Van Voorhis (2016), Bootstrap embedding: An internally consistent fragment-based method, *The Journal of Chemical Physics*, *145*(7), 074,102, doi:10.1063/1.4960986.

Welden, A. R., A. A. Rusakov, and D. Zgid (2016), Exploring connections between statistical mechanics and green's functions for realistic systems: Temperature dependent electronic entropy and internal energy from a self-consistent second-order green's function, *J. Chem. Phys.*, *145*(20), 204,106, doi:10.1063/1.4967449.

Werner, P., and M. Casula (2016), Dynamical screening in correlated electron systems—from lattice models to realistic materials, *Journal of Physics: Condensed Matter*, *28*(38), 383,001, doi:10.1088/0953-8984/28/38/383001.

Yeh, C.-N., S. Iskakov, D. Zgid, and E. Gull (2021), Electron correlations in the cubic paramagnetic perovskite $Sr(V, Mn)o_3$: Results from fully self-consistent self-energy embedding calculations, *Phys. Rev. B*, *103*, 195,149, doi:10.1103/PhysRevB.103.195149.

Yeh, C.-N., A. Shee, Q. Sun, E. Gull, and D. Zgid (2022), Relativistic self-consistent *gw*: Exact two-component formalism with one-electron approximation for solids, doi:10.48550/ARXIV.2202.02252.

Zgid, D., and E. Gull (2017a), Finite temperature quantum embedding theories for correlated systems, *New Journal of Physics*, *19*(2), 023,047, doi:10.1088/1367-2630/aa5d34.

Zgid, D., and E. Gull (2017b), Finite temperature quantum embedding theories for correlated systems, *New Journal of Physics*, *19*(2), 023,047, doi:10.1088/1367-2630/aa5d34.