Title

Quantum Monte Carlo Excited State Orbitals for Optical Gaps of Molecules and Materials

Author

Pineda Flores, Sergio Daniel

Publication Date

2020

Peer reviewed|Thesis/dissertation

Quantum Monte Carlo Excited State Orbitals for Optical Gaps of Molecules and Materials

by

Sergio D. Pineda Flores

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Chemistry

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Professor Eric Neuscamman, Chair
Professor Jeffrey B. Neaton
Professor Martin P. Head-Gordon

Summer 2020

Quantum Monte Carlo Excited State Orbitals for Optical Gaps of Molecules and Materials

Abstract

Quantum Monte Carlo Excited State Orbitals for Optical Gaps of Molecules and Materials

by

Sergio D. Pineda Flores

Doctor of Philosophy in Chemistry

University of California, Berkeley

Professor Eric Neuscamman, Chair

Although amazing progress has been made since the genesis of quantum mechanics in modeling the ground state wave functions and energies of electronic states, modeling the excited electronic states with similar accuracy is still a difficult challenge. In this dissertation orbital optimization of Multi-Slater Jastrow wave functions and its coupling to various Quantum Monte Carlo (QMC) features are used to determine the optical gaps of various systems. The new features being coupled to orbital optimization include: an excited state targeting function, configuration selection for QMC wave functions, a variance matching scheme for optical gaps, and a modified guiding function for sampling within QMC. After a study of the utility of these features, the success of optical gap prediction on both gas phase molecules (aperiodic systems), and condense phase materials (periodic systems) is explored.

For aperiodic systems we found that our QMC optical gap workflow produces predictions on par in terms of accuracy with other standard techniques (e.g. MRCI+Q, CASSCF, CASPT2, EOM-CCSD) for small molecules (e.g. Formaldimine, Thioformaldehyde). In addition, for cases like $\left[C_3N_2O_2H_4Cl\right]^-$ in which state-averaged orbitals (between the ground and first excited state) heavily compromises the accuracy of these states, and multiple reasonable active-space choices lead to very different state energies, our workflow can be advantageous to use.

For periodic systems we found that our QMC optical gap workflow produces predictions on par in terms of accuracy with other standard techniques (e.g. DFT, G_0W_0) for simple bulk materials (e.g. MgO, Trans-Polyaceylene). For more challenging systems, such as bulk transition metal oxides (FeO and MnO) we found that QMC orbital optimization provides the advantage of allowing one to bypass difficult parameterization (e.g. +U value, choice of functional).

This dissertation is dedicated to my mother, Martha Isabel Flores, whose love and support I owe everything to.

Contents

Contents **ii**

List of Figures **iv**

List of Tables **vi**

1 Introduction **1**
1.1 Introduction to the Concept of Electronic Structure Theory 3
1.2 Hartree Fock 4
1.3 Post Hartree Fock Methods 5
1.4 Density Functional Theory 8
1.5 Quantum Monte Carlo 15
1.6 Pseudopotentials 18

2 QMC Specific Theory and Methods **30**
2.1 Excited State Target Function 30
2.2 Modified Guiding Function 31
2.3 Configuration Selection 34
2.4 Variance Matching 35
2.5 Orbital Optimization of Multi-Slater Jastrow Wave Functions 36

3 Aperiodic systems **45**
3.1 Formaldimine (CH_2NH) 45
3.2 Thioformaldehyde (SCH_2) 48
3.3 $\left[C_3N_2O_2H_4Cl\right]^-$ 50

4 Periodic Systems **54**
4.1 Finite size error and explicit variance matching 54
4.2 Magnesium Oxide (MgO) 55
4.3 Trans-Polyacetylene (PAE) 56
4.4 Iron Oxide (FeO) 57
4.5 Manganese Oxide (MnO) 60

4.6 MnO ($\Gamma \to \Gamma$) 65
4.7 MnO ($Z \to \Gamma$) 70

5 Conclusion **75**

Bibliography **77**

A Appendix for Aperiodic Systems **97**
A.1 Computational Details 97
A.2 State Averaging Without Symmetry 99

B Appendix for Periodic Systems **101**
B.1 Computational Details 101

List of Figures

1.1 Aufbau Configuration represented in Slater matrix form and pictorially 5
1.2 The left-most column corresponsd to the Hartree-Fock Aufbau configuration. The center column corresponds to an 'excited' configuration in which an electron from the second lowest orbital is promoted to the fourth lowest. The right-most column illustrates the active space for which electrons and orbitals are considered when producing the linear expansion of the wave function. 7
1.3 Graphical diagram of the workflow for S-CI algorithms. 8
1.4 Key of abbrevations: Ionization Potential = IP, Electron Affinity = EA, Fundamental Gap = E_{Fund}, Optical Gap = E_{Opt}, Binding Energy = E_B, N Particle Ground State = E_N (grd), and N Particle Excited State = E_N $(exct)$ 12
1.5 Sketch of Hedin equation steps (black and blue path) and the GW approximation (black and green path) [28, 27, 76] 14
1.6 Sketch of a 'pseudo wave function' and 'pseudopotential' along with the 'full wave function' and 'full potential'. The $r < r_c$ region refers to the core region [219] . 19
1.7 Sketch of an example ghost state in the production of pseudopotential for Ni. We see a divergence of $D_l(\epsilon)$ indicating a ghost state.[2] 22
1.8 $\langle \eta_l(r)| \, \hat{g}(\epsilon) \, |\eta_l(r)\rangle$ plotted as a function of ϵ. The value of of E_l^{KB} is positive in this case and the black circles correspond to eigenvalues of the total non-local Hamiltoninan. [103] 24

2.1 The objective function Ω (in $(E_\mathrm{h})^{-1}$) during the optimization of the first singlet excited state of SCH_2 using a 6 configuration wave function. Here we compare results using the traditional $|\Psi|^2$ guiding function (A) and our modified guiding function $|\Psi_\mathrm{M}|^2$ (B). The first 25 steps hold the orbitals fixed while optimizing the CI and Jastrow parameters while the last 25 steps optimize the orbitals as well. 33
2.2 A: Energy vs optimization step for H_4 starting from different initial guesses for the orbitals. The orbitals are held fixed and only the CI coefficients and Jastrow variables optimized during the first 25 steps, after which all variables are optimized together. B: Time-step extrapolated DMC energies using the optimized trials wave functions before and after orbital optimization. 44

3.1 Plots of the NLFF fits from Eq. (2.14) (blue curves) for CH_2NH at the 0^o geometry. The excited state variance and energy data points being fit to are shown in red. The arrows demonstrate the variance matching procedure. In the top plot, the arrow takes the ground state variance and uses the NLFF to find the corresponding number of determinants the excited state would need to match that variance. In the bottom plot, the NLFF is used to convert this number into a ground state energy. Note that in this case (for reasons discussed in Section 3.1) we fit the excited state data and interpolate based on an input ground state variance, but the procedure is more typically applied by fitting the ground state data and interpolating based on an excited state variance. 47
3.2 Top row: excitation energies for CH_2NH at various torsion angles. Bottom row: barriers to rotation on the ground state (S0) and excited state (S1) surfaces. Statistical uncertainties were less than 0.04 ev in all cases 48
3.3 The $\left[C_3N_2O_2H_4Cl\right]^-$ anion used in our chlorine-to-π^* charge transfer example. In the ground state, the charge is localized on the (green) chlorine atom, while many of the excited states have the charge distributed in the π network. 50

4.1 Packing arrangement of PAE represented in a 2x2x3 supercell. 57
4.2 VMC (left) / DMC (right) strain curves constructed from various QMC wave functions.[210] The solid lines represent quadratic polynomial fits to the data points to guide the eye. The grey arrow indicates the experimentally determined lattice distortion. 58
4.3 Spin density plots of FeO using various types of QMC wave functions.[210] . . . 59
4.4 The 16 atom super cell (with atoms at the boundary displayed) illustrating the AF2 phase (an antiferromagnetic phase with planes of spin-up and spin-down polarized Mn atoms that alternate in the [111] direction). The purple and gold atoms correspond to spin-up and spin-dn polarized Mn atoms. The red atoms correspond to Oxygen atoms. 65
4.5 A sketch illustrating the high spin arrangement of d-electrons in d-orbitals determined by crystal field splitting of an octahedral geometry and the AF2 phase of MnO. The Mn 3d-orbitals are divided into two neighboring Mn sites (purple or gold) which reside in different ferromagnetic planes (refer to figure 4.4). The relative energy position of d-orbitals on neighboring Mn site is higher than that of the 4s Mn orbital. 67
4.6 The NTO-VMC isosurfaces associated with the particle (Ψ^e) in the conduction band and the hole (Ψ^h) in the valence band. 69
4.7 Comparison of the hole (Ψ^h) in the valence band from various single particle theories. 70

List of Tables

3.1 Excitation energies for the lowest singlet excitation in thioformaldehyde. CASSCF, CASPT2, and MRCI+Q used a full-valence (12e,10o) active space, while SHCI (and the CIPSI calculation from which we generated the MSJ expansion) was performed for all 12 electrons in all orbitals. 49
3.2 Excitation energies in eV for the totally symmetric chlorine-to-π^* transition in $\left[C_3N_2O_2H_4Cl\right]^-$ using BFD effective core potentials and the corresponding VDZ basis set. [29] CASSCF and CASPT2 results are based on either state averaged (SA) orbitals or nearly state specific (SS) orbitals. CASPT2 employed either Roos-Andersson (RA) level shifts [164] or the ionization potential electron affinity (IPEA) approach [69] to deal with intruder states. See Section 3.3 for details. VM-VMC (g,e) stands for Variance Matched VMC with g/e being an integer value representing the number of determinants used in the ground (g) and excited (e) state wave functions. 51

4.1 The optical band gap of MgO determined by different methods. 55
4.2 The optical band gap of PAE determined by different methods. 57
4.3 Table of energies from various QMC wave functions that include the one and two body Jastrow factor and an orbital basis determined from LDA or LDA+U DFT calculations. The parameter values of U = 4.7 eV and J = 0.8 eV values were obtained from constrained DFT calculations determined by Jiang and co-workers [100] 64
4.4 The optical band gap of MnO determined by different methods. 66
4.5 The optical band gap of MnO determined by different methods. 72

Acknowledgments

I would like to thank everyone in the Neuscamman Lab and the Pitzer Center both past and present. Most importantly of these people I'd like to thank my advisor, Eric Neuscamman, who has taught me most of what I know about electronic structure theory, been an example for how to navigate the scientific world, provided guidance in my work, encouraged my scientific curiosity, been respectful of my ideas, and been an all around pleasure to work with. Next I'd like to thank Norman Tubman and Nick Blunt for both really helping me get me started with QMCPACK my first 2 years. I'd also like to thank Nick for being a great friend to hang out with outside of work as well. I'd like to thank Katie Klymko for being very encouraging during my qualifying exam preparations, and whose positive attitude was always infectious. I'd also like to thank Luning Zhao for answering all of my questions about solid state theory but also being fun to hang out with. I'd like to thank Leon Otis for being a great office-mate, and resource for scientific questions I had. Finally, I'd like to thank those I worked with during my time at Sandia Labs, Luke Shulenburger, Joshua Townsend, Raymond Clay and Cody Melton for making me feel welcomed and being fun to work with.

Chapter 1

Introduction

Quantum mechanics (QM) is one of the most important scientific development of the 20th century. Although it's original use was for the esoteric goal of explaining the spectra of atoms, there is hardly an aspect of life today that has not been affected by the 'quantum revolution'. Indeed, nearly every scientific field (chemistry, materials scientist, biology, medicine, computer science etc.) has benefited from QM. Furthermore, the modern economy relies heavily on semiconductor technologies (a technology whose development depends on the understanding of QM). The work of this dissertation is placed in the context of the history of electronic structure theory. Because electronic structure theory can be understood as the applications of QM to the electronic structure of atomic scaled systems, a brief review of some of the key discoveries leading to the first computational self-consistent field calculation in 1951 is worth evaluating.

- 1925 - Wolfgang Pauli proposes the 'Pauli exclusion principle' [147]
- 1925 - Werner Heisenberg, Max Born, Pascual Jordan, invent matrix mechanics [23, 22]
- 1926 - Erwin Schrödinger invents wave mechanics [183]
- 1926 - Wolfgang Pauli derives the hydrogen atom spectrum [146]
- 1927 - Fritz Wolfgang London and Walter Heitler applied QM to the bonding of a hydgrogen molecule [83]
- 1928 - Dirac develops the relativistic wave equation ('Dirac equation') [47]
- 1932 - Linus Pauling publishes what he considers his most important work on "The nature of the chemical bond" [148]
- 1951 - Clemens C. J. Roothaan uses computers to solve the electronic structure of hydrogen and nitrogen molecules [168]

The first electronic structure calculations utilizing computers in 1951 by Clemens C. J. Roothaan was the birth of modern electronic structure theory. The following timeline highlights some of the key advances that have made modern electronic structure techniques (such as post Hartree-Fock methods, Density Functional Theory, and Quantum Monte Carlo) routine.

- 1965 - Walter Kohn and Lu Jeu Sham introduce the Kohn–Sham equations [108]
- 1969 - John Pople pioneers ab initio quantum chemistry methods that use basis sets of either Slater type orbitals or Gaussian orbitals [82]
- 1975 - Anderson develops 'Fixed Node' Diffusion Monte Carlo for fermionic systems [9, 10]
- 1977 - Ceperley, Chester, and Kalos apply Variational Monte Carlo to fermionic system [38]

With these advances the ability to simulate physical processes at the atomic scale to investigate chemical questions became widespread. Instead of simulating simple light weight diatomic molecules, systems with thousands of electrons can now routinely be simulated by a variety of common software packages.[188, 104, 70] In terms of insight from simulation, there is hardly a field of science that has not benefited. Examples of insights achieved include chemical synthetic mechanisms [101], enzyme reactivity in disease [186], and materials discovery [156] just to name a few.

Of all physical processes, this dissertation is most concerned with photoabsorption. Photoabsorption is a light matter interaction that refers to the transition of a system from a ground state to an excited state through the absorption of energy from a photon. The amount of energy required to cause such a transition is referred to as the optical gap (or vertical excitation energy). It is useful to be able to predict the optical gap to interpret spectra and therefore understand processes such as photosynthesis[115], solar cell efficiency[220], and photo-catalytic production of fuels[34, 7].

To produce accurate estimates and insight via simulation, current electronic structure capabilities of optical gap prediction still has many challenges to overcome. Density Functional theory has the notorious self interaction error that cause underestimates of gaps[80]. Active space methods have challenges with scaling and equitable treatment in accuracy of ground and excited states[140]. Perturbation methods have issues of poor zeroth order starting points causing unreliability of predictions.[167, 139]

This dissertation is concerned with both providing accurate estimates and physical insight to the nature of optical gaps through work on the particular class of electronic structure methods referred to as Quantum Monte Carlo (QMC). It is through the coupling of Multi-Slater Jastrow wave function orbital optimization with an excited state targeting function, configuration selection for QMC wave functions, a variance matching scheme for optical gaps, and a modified guiding function for sampling within QMC that we avoid some of the

shortcomings of previously mentioned methods. We will show that orbital optimization of Multi-Slater Jastrow wave functions for QMC when coupled with other QMC features can provide accurate data and insight for optical gaps that is difficult if not impossible to get from other methods.

1.1 Introduction to the Concept of Electronic Structure Theory

Paul Dirac has famously been quoted as saying

> The fundamental laws necessary for the mathematical treatment of a large part of physics and the whole of chemistry are thus completely known, and the difficulty lies only in the fact that application of these laws leads to equations that are too complex to be solved. [46]

The handling of "equations that are too complex to be solved" captures the spirit of the current goals of electronic structure theory community. Electronic structure theory can be understood as describing the motions of electrons of molecular systems. Just as the goal of classical mechanics is to describe the motion of macroscopic particles, the goal of quantum mechanics is to describe the motion of microscopic particles. In particular the described motion is determined in the context of the Born-Oppenheimer approximation[24], an assumption that the motion of electrons and nuclei can be separated due to the different scales of speed of these two types of particles. So ultimately the goal of electronic structure theory is to solve the non-relativistic time-independent Schrödinger (equation 1.1).

$$\hat{H}_{elec} \left|\Psi\right\rangle = E \left|\Psi\right\rangle \tag{1.1}$$

The Hamiltonian term in atomic units takes the form of equation 1.2 for a single particle system, and for a many particle systems it takes on the form of equation 1.3. The 'A' subscrtipt refers to the index of nuclei in the system while the 'i' subscript refers to the index of electrons.

$$\hat{H}_{1\ particle} = -\frac{1}{2}\nabla^2 - \sum_{A}^{M} \frac{Z_A}{r_A} \tag{1.2}$$

$$\hat{H}_{many\ particle} = -\sum_{i} \frac{1}{2}\nabla_i^2 - \sum_{i}^{N}\sum_{A}^{M} \frac{Z_A}{r_{iA}} + \sum_{i}^{N}\sum_{j>i}^{N} \frac{1}{r_{ij}} \tag{1.3}$$

While the Schrödinger equation of a single particle (equation 1.2) such as a hydrogen atom can be exactly solved, any system that includes more than one particle (equation 1.3) is classified as a many-body problem and does not have an analytic solution. The source of

the trouble is the third term in equation 1.3 which corresponds to the instantaneous Coulomb repulsion between pairs of electrons.

Quantum chemist and solid sate physicists alike endeavor to find computational tractable solutions to the many-body problem. As a result, there are a myriad of electronic structure techniques that approach the problem from a variety of perspectives. Some focus on reducing the scaling of solutions that are approximately exact, others focus on "cooking in" physical insight to the technique that approach approximately exact solutions.

In the sections that follow we briefly review the electronic structure techniques relevant to the work in this dissertation. The techniques reviewed are: 1) the Hartree Fock method, 2) post Hartree Fock methods, 3) Density Fucntional Theory 4) Quantum Monte Carlo methods, and 5) pseudopotentials.

1.2 Hartree Fock

We review Hartree-Fock (HF) because it is often a starting point for modern Quantum Chemistry methods and a brief overview will allow us to define some terms that will be used frequently throughout this dissertation. In terms of HF's relevance to QMC methods, often the HF orbitals are chosen as the Single Particle Orbital set (SPO set) to form a trail wave function.

Overview

The Hartree Fock method uses the determinant of a Slater matrix as the ansatz (trial wave function) to solve the time-independent Schrödinger equation. The Slater matrix (figure 1.1) is made up of N spin orbitals ($\phi_\sigma(r)$) in which 'r' corresponds to position coordinates and 'σ' corresponds to spin. Due to the variational principle (which states that any trial wave function will produce an upper bound to the true ground state energy [200]), a solution is found by simply varying the orbitals until a minimum of the total energy is found. Using Lagrange's method of undetermined multipliers on a functional of the energy constrained with orthonormal orbitals (equation 1.4 [200] in which ϵ_{ba} are the Lagrange multipliers) leads to the Hartree-Fock equations (equation 1.5 in which kets always represent electron 1).[200]

$$\mathcal{L}[\{\phi_a\}] = E_0[\{\phi_a\}] - \sum_{a=1}^{N}\sum_{b=1}^{N}\epsilon_{ba}(\langle a|b\rangle] - \delta_{ab}) \tag{1.4}$$

$$[h(1) + \sum_{b=1}^{N} J_b(1) - K_b(1)]\,|\phi_a\rangle = \epsilon_a\,|\phi_a\rangle \qquad a = 1, 2, ..., N \tag{1.5}$$

The terms in the brackets of equation 1.5 is known as the Fock operator while the Hartree Fock potential (v_{hf}) refers to $\sum_{b=1}^{N} J_b(1) - K_b(1)$ in which $J_b(1)\,|\phi_a\rangle = [\int dx_2 \phi_b^*(2) r_{12}^{-1} \phi_b(2)]\,|\phi_a\rangle$ and $K_b(1)\,|\phi_a\rangle = [\int dx_2 \phi_b^*(2) r_{12}^{-1} \phi_a(2)]\,|\phi_b\rangle$. The interpretation of equation 1.5 is that the

orbitals each experience a effective potential and so explicit coulomb electron-electron correlation is missing from the theory.

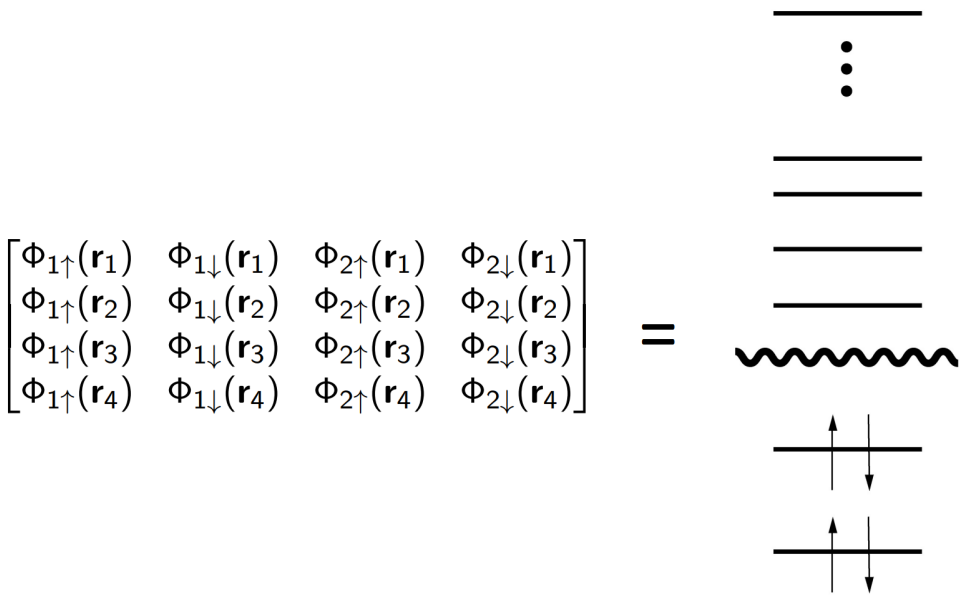


Figure 1.1: Aufbau Configuration represented in Slater matrix form and pictorially

The electron-electron correlation missing from Hartree Fock is further divided into weak and strong correlation (also known as dynamic and static correlation respectively). Weakly correlated systems are those in which a single configuration mean field method like Hartree Fock does an adequate job. For these weakly correlated systems perturbation theory can adequately recover missing effects. Strongly correlated systems are those in which a single configuration mean field method does not even qualitatively describe the system correctly. Trying to use a single configuration as a zeroth order approximation for strongly correlated systems results in perturbation theory failing disastrously.

1.3 Post Hartree Fock Methods

The goal of post Hartree Fock methods (also referred to simply as wave function based methods in this dissertation) is to capture the missing correlation effects absent in Hartree Fock. Typically this missing correlation is captured through a linear expansion of the trial wave function in terms of 'configurations' (which QMC methods can use as the starting point for a trial wave function). In this dissertation we take advantage of the wave function active

space methods Configuration Interaction (CI) [109, 57, 137, 157], Selective Configuration Interaction (S-CI) [93, 189, 182, 71, 33, 169, 214] and Complete Active Space SCF method (CASSCF) [166, 141, 99, 117] in order to select configurations for our QMC methods.

Overview

Most wave function methods begin with the Hartree-Fock solution (the optimal orbitals for a single slater wave function). When speaking of the basis for these methods one could divide it into the single-particle orbital set (SPO set) and the many-body basis ('exctied determinants'/configuartions with respect to the Aufbau configuration constructed from Hartree Fock orbitals). If all SPO and excited configurations are taken into account (referred to as 'full configuration interaction') then the exact solution (in the Fock space defined by the SPO set) can be determined.[84, 200] But the number of configurations necessary for a calculation containing K orbitals and N electrons is $\binom{K}{N} = \frac{K!}{N!(K-N)!}$, so FCI scales exponentially [84, 200] and only works for small systems in practice. Because FCI is too expensive some sort of approximation to the basis must be made, either what configurations to include or how to determine the weight of these configurations. In this way a trade off will always be made between affordability (whether computational resources or human wall time) and accuracy (number of configurations to consider and method of approximating their weight in the expansion of the wave function).

For systems that require multi-reference description of states (such as the the larger systems mentioned earlier) the CASSCF method is often used. The idea of the method is to choose a subset of electrons and orbitals to perform FCI to hopefully capture the multi-reference description.[84] Then there is an additional step of allowing the orbitals to relax in the presence of the new multi-slater expansion.[84] Although this method will help describe the multi-reference state the draw backs are that human input in necessary to determine the active space and dynamic correlation is generally missing. But a second order perturbation calculation from the CASSCF reference (known as CASPT2) helps recover the dynamic correlation.[84] Yet CASPT2 can suffer from 'intruder states' similar to how MP2 does.[165, 40] Simply increasing active space to attempt to avoid this issue is not always possible and any type of "removal" of intruder states (through 'shifts' or otherwise) inevitably introduces a new possibility of errors.[229]

For Selective-Configuration Interaction the idea is to use perturbation theory repeatedly to select what configurations to include in the linear expansion of the wave function.[93] The schematic of the algorithm is presented in figure 1.3. In summary the steps to the algorithm are:

1. start with the Hartree-Fock Aufbau configuration,
2. perform a first order perturbation correction to the wave function to have candidate configurations to include in the linear expansion of the wave function,
3. use a threshold weight to decide which configurations to keep in your linear expansion,

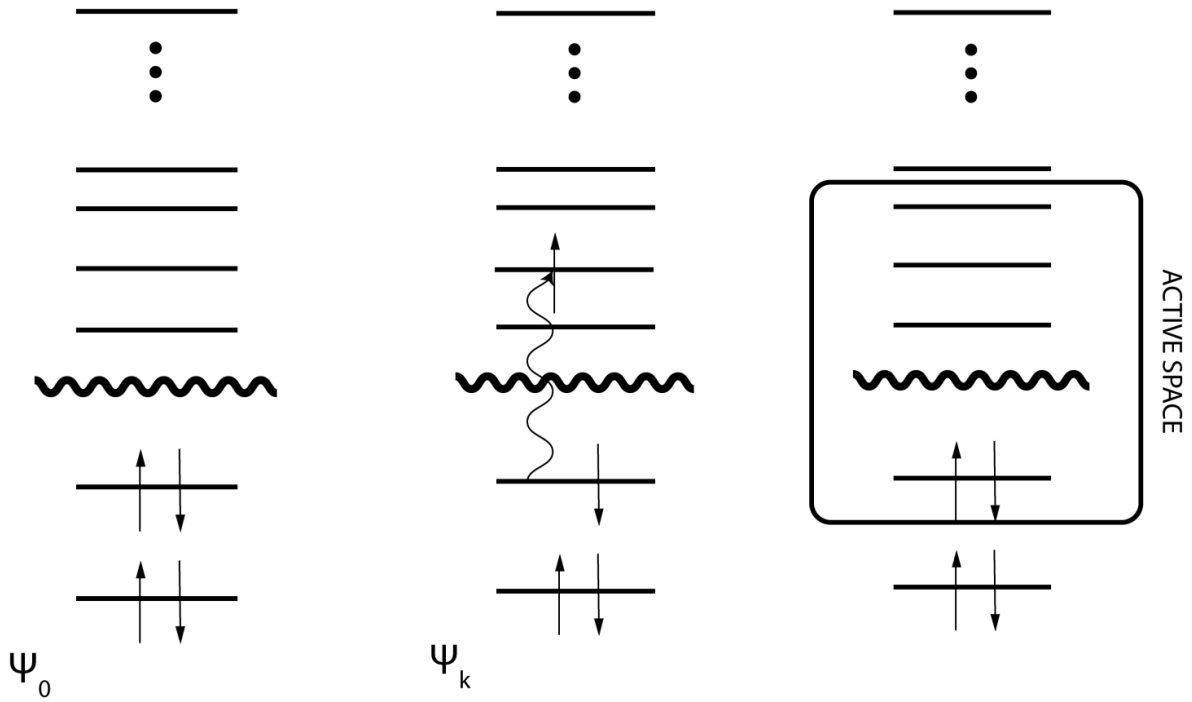


Figure 1.2: The left-most column corresponsd to the Hartree-Fock Aufbau configuration. The center column corresponds to an 'excited' configuration in which an electron from the second lowest orbital is promoted to the fourth lowest. The right-most column illustrates the active space for which electrons and orbitals are considered when producing the linear expansion of the wave function.

4. diagonalize the Hamiltonian in the linear expansion you have

5. repeat steps 2-4 until energy converges or some limit of configurations is reached.

Although S-CI allows one to consider a larger number of electrons and orbitals for the active space it is ultimately limited once again by the exponential rise of necessary configurations as system size grows and has no guarantee in treating the ground/excited state fairly with respect to each other.

A comment on the practical limitation of the size of active spaces depends on the ever changing current state of the art in algorithm development for currently available hardware. Recently, NVIDA GPU programming of CASSCF has managed to perform routine active space calculations of 18 electrons in 18 orbitals (2.4 billion determinants) [56]. Popular software (Molpro, Gamess, etc.) employing active space methods is typically limited to electrons and orbitals in the teens.[217, 74] For S-CI methods the active space can be larger, such as

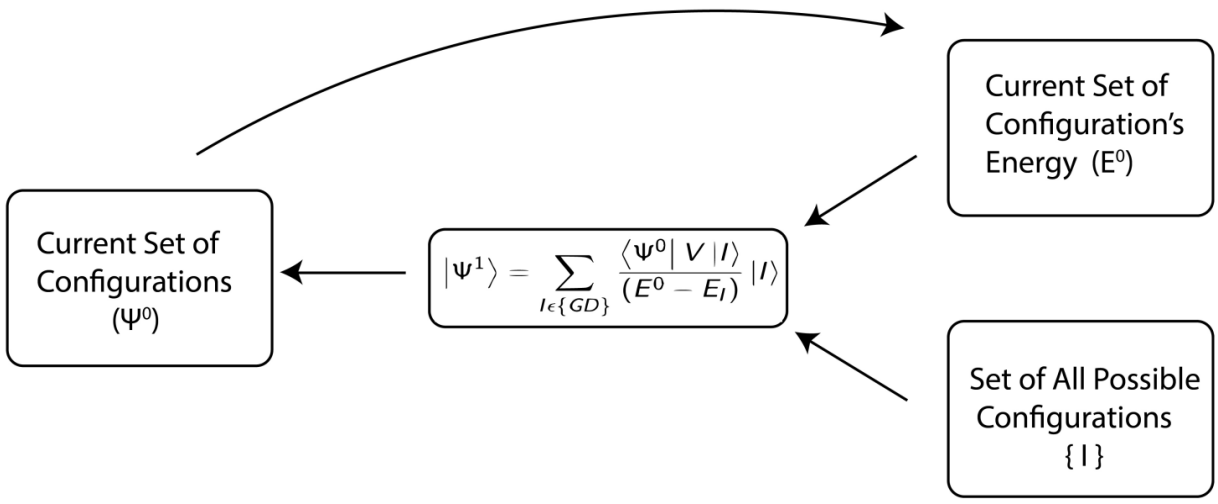


Figure 1.3: Graphical diagram of the workflow for S-CI algorithms.

the 'fast semistochastic heat-bath configuration interaction' ability to perform calculations on 28 electrons in 76 orbitals[114] or even 44 electrons in 44 orbitals [198].

In addition to practical active space size limitations of any truncated CI method, these methods typically suffer from not being 'size consistent'.[13, 84, 200], Size consistency refers to the property that for systems consisting of multiple non-interacting subunits, the calculated energy of the entire system equals the sum of the calculated energies of the subunits. For a system of two non-interacting subunits (A and B) that means E(A+B) = E(A)+E(B) will be true for size consistent methods.[84, 200] The typical lack of size consistency for truncated CI methods is problematic, but especially so for solids when sampling the Brillouin zone. It is these active space size limits and lack of size consistency that makes active space based methods mostly limited to atoms or small molecules, while periodic systems tend to utilize DFT and GW methods.

1.4 Density Functional Theory

In the process of reviewing the basics of Density Functional Theory (DFT) we will also define the fundamental/optical gap, and discuss how well DFT performs in predicting these. With respect to QMC, DFT serves as another source of SPO sets.

Overview

DFT is the most popular electronic structure theory method for determining quantum mechanical properties of atomic-scale systems. It reduces the many-body problem of solving

the time-independent Schrödinger equation for an electronic wave function $\Psi(\mathbf{r})$ (a function of 3N dimensions, in which N is the number particles) to that of finding the optimal electronic density $n(r)$ (a function of only 3 dimensions).[145] It is through the use of functionals – functions of functions – that the total energy of the system can be determined via the electronic density (equation 1.6).

$$E[n] = T[n] + U[n] + \int V(r)n(r)d^3r \tag{1.6}$$

DFT is a formally exact theory, proven by Hohenberg and Kohn [89], in the sense that given the exact ground state density $n(r)$ there exists a universal functional ($F[n] = T[n] + U[n]$) that can be used to evaluate equation 1.6 to yield the exact ground state energy. Any shortcoming of DFT in practice is due to the approximations made for these functionals, since their general form is not known. Furthermore, no general prescription (that avoids an explicit use of many body wave functions) for how to determine $n(r)$ for a particular system is known. It is with the existence of the Hohenberg–Kohn (HK) theorems [89] that DFT is made possible in practice. These theorems state that: (1) there is a one-to-one mapping between a given external potential (i.e. the system of interest) to its $n(r)$, and (2) the ground state $n(r)$ is that which minimizes the total energy functional. With these two theorems the ground state $n(r)$ is determined *uniquely* by varying the density until the total energy is minimized; all that is left is to make approximations for $T[n]$ and $U[n]$.

The Kohn–Sham (KS) equations [108] address how to make systematic approximations to $T[n]$. By mapping the interacting system to a fictitious non-interacting system (represented by the Kohn-Sham determinant) in which particles experience an effective potential that captures all exchange-correlation effects, the $T[n]$ can be approximated by its non-interacting form $T_s[n]$ (equation 1.7).

$$T_s[n] = \sum_{i=1}^{N} \int d\mathbf{r}\, \phi_i^*(\mathbf{r}) \left(-\frac{\hbar^2}{2m}\nabla^2\right) \phi_i(\mathbf{r}) \tag{1.7}$$

With the kinetic energy approximated, $U[n]$ needs to capture all other quantum mechanical effects, including exchange-correlation effects and any further differences $T[n] - T_s[n]$.

Functionals

The exchange-correlation (XC) functional is known exactly only for the free-electron gas, but approximations that work well in practice have been made. One of the earliest being the LDA approximation (equation 1.8) [37, 152, 72], derived from the homogeneous electron gas (HEG) by Perdew and Wang. To determine the $\varepsilon_{\mathrm{XC}}(n)$ used for LDA, Perdew and Wang fit a function [152] to accurate QMC calculations (produced by Ceperly[37]) of the HEG at various densities.

$$E_{\mathrm{XC}}^{\mathrm{LDA}}[n] = \int \varepsilon_{\mathrm{XC}}(n)n(\mathbf{r})\, \mathrm{d}^3\mathbf{r} \tag{1.8}$$

But because LDA is constructed under the assumption of uniform density, taking into account the non-homogeneity of the true density is the natural next progression in the design of the XC-functional. These XC-functional are referred to as Generalized Gradient Approximations (GGA) and have the general form of equation 1.9.

$$E_{\mathrm{XC}}^{\mathrm{GGA}}[n_\uparrow, n_\downarrow] = \int \varepsilon_{\mathrm{XC}}(n_\uparrow, n_\downarrow, \nabla n_\uparrow, \nabla n_\downarrow) n(\mathbf{r})\, \mathrm{d}^3\mathbf{r} \tag{1.9}$$

An issue with local functionals (LDA and GGA) is that they suffer from self-interaction error (SIE), which can most easily be illustrated by considering single electron systems. A single electron should not Coulombically interact with itself yet the use of local functionals will yield a non-zero interaction energy. Hartree-Fock does not suffer from SIE due to its exchange energy exactly canceling the self coulomb interaction. Another physical picture illustrating the short comings of local functionals is the potential experienced by an electron as it moves away from a finite system that is neutral. The electron should be interacting with it's hole in a way that scales as $\lim_{|r|\to\infty} v_{xc}(r) = -\frac{1}{|r|}$ (proven by Almbladh and Barth.[8]) but this is not the case for local functionals [8]. These observation suggests that to improve the XC-functional the exact exchange of Hartree-Fock should be incorporated into XC-functionals[15] which led to the rise of 'hybrid functionals' (such as PBE0, equation 1.10 [5]).

$$E_{\mathrm{xc}}^{\mathrm{PBE0}} = \frac{1}{4} E_{\mathrm{x}}^{\mathrm{HF}} + \frac{3}{4} E_{\mathrm{x}}^{\mathrm{PBE}} + E_{\mathrm{c}}^{\mathrm{PBE}} \tag{1.10}$$

Although hybrid functionals perform better than local functionals there are still SIE present because only a fraction of the correct asymptotic behavior is recovered for hybrid functionals.[8] To more adequately correct the long range potential behavior 'range separated functionals' were develop. The correct asymptotic potential is recovered by partitioning the electron-electron interaction into a short range (typically consisting of local functional exchange) and long range (typically consisting of HF exchange). The HSE06 functional (equation 1.11 [86, 85]) has a division of short and long range potentials but numerical tests led the developers to use HF for short range exchange potentials and PBE for long range.[87]

$$E_{\mathrm{xc}}^{\omega\mathrm{PBEh}} = a E_{\mathrm{x}}^{\mathrm{HF,SR}}(\omega) + (1-a) E_{\mathrm{x}}^{\mathrm{PBE,SR}}(\omega) + E_{\mathrm{x}}^{\mathrm{PBE,LR}}(\omega) + E_{\mathrm{c}}^{\mathrm{PBE}} \tag{1.11}$$

DFT Band Gap

Before we assess DFT band gap predictions we define the ionization potential (IP), electron affinity (EA), Fundamental Gap (E_{Fund}) and optical gap (E_{Opt}).

$$IP = E_{N-1} - E_N \tag{1.12}$$

$$EA = E_N - E_{N+1} \tag{1.13}$$

$$E_{\mathrm{Fund}} = (E_{N+1} - E_N) - (E_N - E_{N-1}) = IP - EA \tag{1.14}$$

The fundamental gap is measured experimentally via photoemission and inverse-photoemission experiments that calculate the IP and EA energies respectively. Both of these experiments involve a change in the number of particles. In contrast the optical gap is the difference in energy between the ground state and first excited state with the same number of particles and is measured experimentally via absorption spectroscopy of a dipole allowed transition. The exciton binding energy refers to the Coulomb interaction between the electon and it's hole in the excited state. A graphical summary of all these definitions is found in figure 1.4.

To calculate the fundamental gap two ground state calculations of a system containing N and N+1 particles must be done. But often the difference in orbital energies $\epsilon_{LUMO} - \epsilon_{HOMO}$ of an N particle calculation is used to approximate the fundamental gap. This approximation is appropriate for HF due to Koopman's theorem which states that IP/EA can be approximated as $IP = -\epsilon_{HOMO}$ and $EA = -\epsilon_{LUMO}$ [200] but the orbital energies from DFT do not have physical meaning (with the exception of the HOMO being $-IP$).[145] Therefore an approximation of the fundamental gap via calculating $\epsilon_{LUMO} - \epsilon_{HOMO}$ using DFT orbital energies have no reason to correspond to the gap. Perdew and coworkers proved that the fundamental gap is calculated as equation 1.15.[154]

$$\epsilon(N+1)_{HOMO} - \epsilon(N)_{HOMO} + C \tag{1.15}$$

The term $\epsilon(N+1)_{HOMO}$ refers to the DFT HOMO orbital energy of a N+1 particle system, the term $\epsilon(N)_{HOMO}$ refers to the DFT HOMO orbital energy of a N particle system and the 'C' term refers to the derivative discontinuity of the energy with respect to particle count evaluated at N particles. So while a single band structure calculation containing N particles can estimate $\epsilon(N+1)_{HOMO} - \epsilon(N)_{HOMO}$ as $\epsilon(N)_{LUMO} - \epsilon(N)_{HOMO}$ the Kohn-Sham band structure does not include the finite and positive derivative discontinuity of the Exchange-Correlation energy as a function of number of particles. Therefore even if the exact XC-functional was being used for the single band structure calculation containing N particles the correct fundamental band gap will not be predicted. [153, 151, 187, 172]

Despite $\epsilon(N)_{LUMO} - \epsilon(N)_{HOMO}$ using DFT energies not having a rigorous justification for estimating the fundamental gap it is commonly done and has proven accurate quite often.[173, 134, 221] But it has been shown that this accuracy is rather fortuitous. [98, 128, 176]

Many-Body Perturbation Theory

Results from G_0W_0 calculations are considered when comparing the results of our QMC work to other state of the art methods. Therefore, in this section we review the basics of Many-Body theory necessary to have some basic understanding and ability to interpret results. We follow notation conventions commonly found in standard Many-Body theory textbooks and online tutorials provided by software packages. [3, 60, 121, 76]

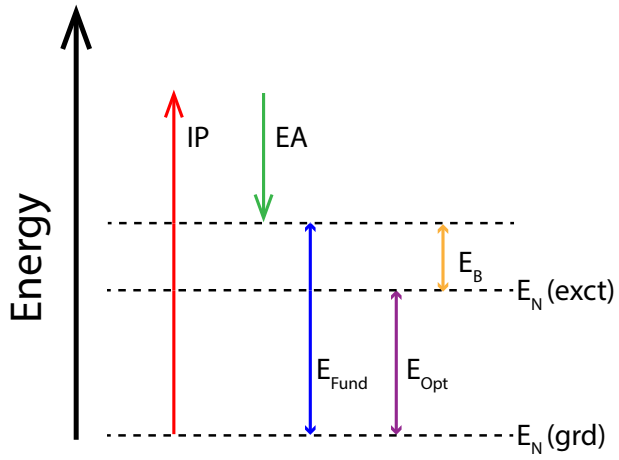


Figure 1.4: Key of abbrevations: Ionization Potential = IP, Electron Affinity = EA, Fundamental Gap = E_{Fund}, Optical Gap = E_{Opt}, Binding Energy = E_B, N Particle Ground State = $E_N\ (grd)$, and N Particle Excited State = $E_N\ (exct)$

Green's function and self energy

Any single-particle operator can be evaluated once the Green's function is known so we begin with the time-ordered Green's function $G(1,2)$. The integer coordinates represent both space and time, the $\hat{T}[...]$ term is the time ordering operator of the creation and annihilation field operators and Θ_0^N represents the N particle state. The field operators create/destroy a particle that is an eigenfunction of the coordinate operator, and in equation 1.16 are in the Heisenberg representation. An intuitive interpretation of the Green's function is that it determines the probability of a particle moving from one point to another in a given amount of time which is why it is sometimes referred to as the propagator.

$$G(1,2) = -i \left\langle \Theta_0^N \left| \hat{T}\left[\hat{\psi}(1)\hat{\psi}^\dagger(2)\right] \right| \Theta_0^N \right\rangle \tag{1.16}$$

To see how single excitation information can be extracted from Green's function it is useful to re-express it in the Lehmann representation (equation 1.18) with the Lehmann amplitudes (or Feynman-Dyson amplitudes) defined as equation 1.17. For the Lehmann amplitudes μ represents the chemical potential of the system (or HOMO of the N particle system), the ϵ terms are the single excitation energies of the system and the 'i' subscript refers to the set of quantum numbers associated with the N+1 or N-1 states. Note from equation 1.18 that if one locates the poles of the Green's function in the complex plane then information of the absorption spectrum is secured.

$$\Psi_i(r) \equiv \begin{cases} \langle \Theta_0^N | \hat{\psi}(r) | \Theta_i^{N+1} \rangle & \epsilon_i > \mu \\ \\ \langle \Theta_i^{N-1} | \hat{\psi}(r) | \Theta_0^N \rangle & \epsilon_i < \mu \end{cases} \tag{1.17}$$

$$G(r_1, r_2; \omega) = \sum_i \frac{\Psi_i(r_1)\Psi_i^\dagger(r_2)}{\omega - \epsilon_i + i\eta\, sign(\epsilon_i - \mu)} \qquad \eta \to 0^+ \tag{1.18}$$

So now that we have established that we want the full interacting Green's function we introduce the Dyson equation (equation 1.19) which describes the relationship between the non-interacting Green's functions (G_0) and fully interacting Green's functions via the self-energy (Σ).

$$G(12) = G_0(12) + \int G_0(13)\, \Sigma(34)\, G(42) d34 \tag{1.19}$$

The self-energy describes the screened Coulomb interaction of electrons and can be used to solve for the quasi particle excitation energies.

$$\Big[\hat{h}_0(r_1) + v_H(r_1)\Big]\Psi(r_1) + \int \Sigma(r_1, r_2; \epsilon^{QP})\Psi(r_2) dr_2 = \epsilon^{QP}\Psi(r_1) \tag{1.20}$$

At this point the main goal is to determine Σ.

Hedin's equation

In 1965 Hedin derived a way to compute the exact Green's function[81] which involved computing the irreducible polarizability (χ), the dynamically screened interaction (W) and the vertex function. The process is summarized by the black and blue path in figure 1.5. In principle the equations should be solved self-consistently, but in practice this is difficult to do. This leads to the single shot GW approximation.

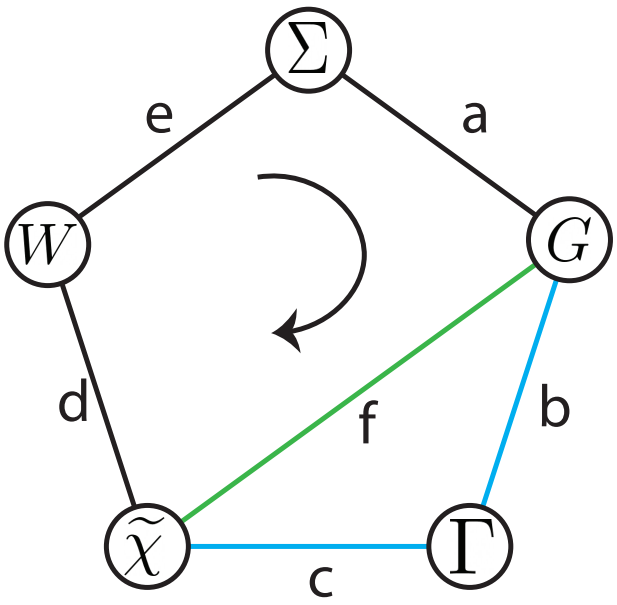
The GW appoximation

Because solving Hedin's equations is very demanding an approximate solution referred to as the 'GW approximation' approximates the vertex function as equation 1.21

$$\Gamma(12; 3) \approx \delta(r_1 - r_2)\, \delta(t_1 - t_2)\delta(r_1 - r_3)\, \delta(t_1 - t_3) \equiv \Gamma^{GW}(12; 3). \tag{1.21}$$

This leads to simplification of the necessary equations for Σ (summarized graphically as the black and green path of figure 1.5). Iterating through these steps once for Σ is referred to as G_0W_0. Continuing through the cycle updating both G and W is the 'fully self consistent GW', and updating just G is referred to as the partially self consistent version (GW_0).

To begin the GW cycle an initialization for the self-energy is needed. The Kohn-Sham exchange correlation potential is used as an approximation to the self-energy, the Kohn-Sham states are approximations to the quasiparticle states and the Kohn-Sham eigenvalues



$$a : G = G_0 + G_0 \Sigma G$$
$$b : \Gamma = 1 + (\delta\Sigma/\delta G) G G \Gamma$$
$$c : \widetilde{\chi} = -iGG\Gamma$$
$$d : W = v + v\widetilde{\chi}W$$
$$e : \Sigma = iGW\Gamma$$
$$f : \widetilde{\chi} = -iGG$$

Figure 1.5: Sketch of Hedin equation steps (black and blue path) and the GW approximation (black and green path) [28, 27, 76]

approximations to the quasiparticle energies. In this way the quasiparticle energies can be solved using first order perturbation theory via equations 1.22 and 1.23 .[95, 94]

$$\epsilon^{QP} = \epsilon^{KS} + Z\langle \Psi^{KS} | \Sigma(\epsilon^{KS}) - v_{xc} | \Psi^{KS} \rangle \tag{1.22}$$

$$Z \equiv \left[1 - \langle \Psi^{KS} | \frac{\Sigma(\epsilon)}{\epsilon^{KS}} | \Psi^{KS} \rangle \right]^{-1} \tag{1.23}$$

From the presented description it is clear that the accuracy of a G_0W_0 calculation will depend on the quality of initial approximation being used. Because the initial approximation comes from a DFT calculation the Kohn-Sham energies and SPOs are the limiting factor of G_0W_0 accuracy and need to be near the true quasiparticle energy and states. For simple s-p bonded systems in which DFT generally can describe the system well, G_0W_0 can easily correct the typical underestimation of the gap. [116] But ultimately for systems that are difficult for DFT to describe correctly, such as transition metal oxides, G_0W_0 becomes less reliable. [174, 100, 110] One may ask if there is a way to assess the quality of the input SPO set for G_0W_0. One attempt has been matching the HOMO to the experimental ionization potential as a way to improve the initial quasiparticle energy and mitigate the SIE in the creation of the quasiparticle states. [158, 59] Later in this dissertation we describe how QMC can be used to assess the quality of SPO sets for G_0W_0.

1.5 Quantum Monte Carlo

Quantum Monte Carlo is a method that: 1) utilizes the full interacting Hamiltonian and 2) evaluates integrals using Monte Carlo stochastic sampling. The consequence of point 1) suggest that one can interpret its use as tackling the strong correlation problem in a 'variational and head on manner' as the other wave function methods we mentioned do.[14, 206, 63, 79] But the main advantage of using QMC methods lies in point 2) because the trial wave function can take on forms that may not be convenient for traditional wave function methods.[14, 206, 63, 79]

In the following subsections we briefly review the basics of the QMC methods (Variational Monte Carlo and Diffusion Monte Carlo) but will be presenting more specific details relevant to the work in this dissertation later. The reader is encourage to read more comprehensive reviews of the methods. [206, 63, 79]

Variational Monte Carlo

The wave function methods of the previous section are limited to a Gaussian basis due to the reliance of analytic formulas for evaluating integrals [6]. The main advantage of VMC is that integrals are evaluated via Monte Carlo stochastic sampling therefore the form of the wave function is much more flexible.[14, 206, 63] This allows for the wave function (wfn) to take on many different forms such as the Slater-Jastrow wfn [120, 208, 216], Jastrow correlated antisymmetrized geminal power (JAGP) wfn [138, 36], Multi-Slater Jastrow wfn[62, 178, 179], and backflow transformation [61, 131, 135]. We will speak more about the Jastrow factor in section 1.5.

VMC uses the idea that the multidimensional integrals needed to evaluate the expectation values of the energy of a system can be determined using the Monte Carlo integration technique.[14, 206, 63] Equations 1.24 through 1.27 summarize how the Monte Carlo technique is used to evaluate the energy. Equation 1.24 is the expectation value of the energy of a trial wave function re-expressed as the integration of the 'local energy' ($E_L(\boldsymbol{R})$ equation 1.26) according to the probability distribution of the wave function ($\rho(\boldsymbol{R})$ equation 1.25). Equation 1.27 summarizes how the integration can then be estimated via sampling the local energy function M times according to the distribution of $\rho(\boldsymbol{R})$ using the Metropolis-Hasting algorithm[126].

$$E_v = \frac{\int \Psi^*(\boldsymbol{R})\hat{H}\Psi(\boldsymbol{R})d\boldsymbol{R}}{\int \Psi^*(\boldsymbol{R})\Psi(\boldsymbol{R})d\boldsymbol{R}} = \int d\boldsymbol{R}\ \rho(\boldsymbol{R})E_L(\boldsymbol{R}) \tag{1.24}$$

$$\rho(\boldsymbol{R}) = \frac{|\Psi_T(\boldsymbol{R})|^2}{\int |\Psi_T(\boldsymbol{R})|^2 d\boldsymbol{R}} \tag{1.25}$$

$$E_L(\boldsymbol{R}) = \frac{\hat{H}\Psi_T(\boldsymbol{R})}{\Psi_T(\boldsymbol{R})} \tag{1.26}$$

$$E_v \approx \frac{1}{M}\sum_{m=1}^{M} E_L(\boldsymbol{R}_m) \tag{1.27}$$

Estimated VMC energies have statistical uncertainty that is determined using standard statistical techniques.[58] If we consider an estimate of E_v from $M1$ samples (referred to as E_1 with a variance of σ_1^2), another from $M2$ samples (E_2 / σ_2^2) and so on ... These values of E_X and σ_X^2 are independent identically distributed and so the familiar findings of the central limit theorem can be used. As sampling approaches infinity the distribution of the E_X approaches a normal distribution with a mean value of the true E_v value and a variance σ^2/M. Note the feature that the statistical uncertainly falls as $1/\sqrt{M}$ regardless of the dimensionality of the problem, in contrast to deterministic numerical integration techniques such as the Simpson's integration method [63].

One last property of VMC that should be noted is the "zero-variance" property.[14, 206, 63] As illustrated above, how many samples must be taken in order to reduce the error to the desirable amount depends in part on the variance of the local energy. As the trial wave function approaches the exact wave function the variance approaches zero. Therefore optimizing the parameters of a wave function improves the energy, and the statistical uncertainty. As will be discussed later, optimizing certain parameters (such as the orbital rotation parameters) also improves the quality of the nodal surface of the wave function and can therefore improve the Diffusion Monte Carlo estimate of the energy. [14, 206, 63]

Trial Wave Function Form

To describe the ground and excited states in this dissertation we will use a wave function of the Multi-Slater Jastrow (MSJ) form. The MSJ wave function can be written as

$$\Psi(\vec{r}) = \Phi(\vec{r})\,\psi_J(\vec{r}) \tag{1.28}$$

$$\Phi(\vec{r}) = \sum_{I=0}^{N} c_I D_I(\vec{r}) \tag{1.29}$$

in which ψ_J is the Jastrow factor and Φ is the linear combination of Slater determinants D_I.

Although there are a wide variety of forms for the Jastrow factor Ψ_J, for this study we employed only the one-, two-, and three-body Jastrows.

$$\begin{aligned}\psi_J =& \exp\Big(u_1(r_{\sigma i}) + \frac{1}{2}\sum_{\sigma\sigma' ij} u_2(r_{\sigma i}, r_{\sigma' j}) \\ &+ \frac{1}{2}\sum_{I}\sum_{\sigma\sigma' ij} u_3(r_{\sigma i}, r_{\sigma' j}, R_I)\Big)\end{aligned} \tag{1.30}$$

$$u_1(r_{\sigma i}) = \sum_{\sigma i}\sum_{I} u_{\sigma I}(|r_{\sigma i} - R_I|) \tag{1.31}$$

$$u_2(r_{\sigma i}, r_{\sigma' j}) = u_{\sigma\sigma'}(|r_{\sigma i} - r_{\sigma' j}|) \tag{1.32}$$

$$u_3(r_{\sigma i}, r_{\sigma' j}, R_I) = \\ u_{\sigma\sigma' I}(|r_{\sigma i} - R_I|, |r_{\sigma' j} - R_I|, |r_{\sigma i} - r_{\sigma' j}|) \tag{1.33}$$

Our Jastrow takes the form of equations 1.30-1.33 in which indices I are for ions, i/j are for electrons, and σ/σ' are for spins. The function $u_{\sigma\sigma'}$ takes on one of two forms, u_{ss} or u_{os}, depending on whether electrons σ and σ' have the same spin or opposite spins, and these two forms are constructed using splines [104] so as to guarantee that the appropriate electron-electron cusp conditions are satisfied[102]. The functions $u_{\sigma I}$ and $u_{\sigma\sigma' I}$ are similarly constructed of splines.[104] The function $u_{\sigma I}$ can either be formulated to enforce the nuclear cusp condition or to be cusp-free in cases where either a pseudopotential is used or the cusp is built into the orbitals.[63] The function $u_{\sigma\sigma' I}$ take el-el, el-ion, el-el-ion separations as arguments and is identical to that proposed by Drummond and coworkers.[49] Because the Jastrow factor reproduces the cusps determined by interparticle distances the multi-Slater expansion can be much smaller than would otherwise be needed.

Diffusion Monte Carlo

Diffusion Monte Carlo (DMC) is a projection technique that allows one to sample from the exact wave function (Ψ_0) distribution and therefore compute the exact energy from a mixed expectation value.[14, 206, 63, 79, 9, 10]

$$E_0 = \frac{\langle \Psi_0 | \hat{H} | \Psi \rangle}{\langle \Psi_0 | \Psi \rangle} \tag{1.34}$$

Note that the calculation of $\hat{H}$ does not need to be evaluated on Ψ_0, instead the usual local energy according to trial wave function is evaluated but sampling occurs according the mixed distribution of $f(\boldsymbol{R}) = \Psi_0(\boldsymbol{R})\Psi(\boldsymbol{R}) / \int d\boldsymbol{R} \Psi_0(\boldsymbol{R})\Psi(\boldsymbol{R})$ instead of the usual $\rho(\boldsymbol{R})$. To begin, the time-dependent Schrödinger equation is rewritten in it's imaginary time form ($\tau \to -it$)

$$|\Psi(t)\rangle = e^{-(\hat{H} - E_T)t} |\Psi\rangle \tag{1.35}$$

in which E_T is treated as some parameter that can be chosen. The spectral decomposition of $e^{-(\hat{H}-E_T)t}$ can then be inserted to equation 1.35 and taken to limit of infinite time $\lim_{\tau\to\infty} |\Psi(t)\rangle = \lim_{\tau\to\infty} e^{-(E_0 - E_T)\tau} |\Psi_0\rangle \langle \Psi_0 | \Psi \rangle$ resulting in equation 1.36 .

$$\lim_{\tau\to\infty} |\Psi(t)\rangle \approx |\Psi_0\rangle \tag{1.36}$$

Now if we re-write our current expression (equation 1.35) in position basis with the defined Green's function (equation 1.38) the result is equation 1.37. In practice the Green's function is determined in the limit of short time using the Trotter-Suzuki formula 1.39. Now with the the similarity transformed Green's function (equation 1.40) we have a way of evaluating the mixed distribution $f(\boldsymbol{R})$ we originally wanted via equation 1.41.

$$\Psi(R_f, t) = \int dR_i G(R_f|R_i; t)\Psi(R_i) \tag{1.37}$$

$$G(R_f|R_i; t) = \langle R_f| \, e^{-(\hat{H}-E_T)\tau} \, |R_i\rangle \tag{1.38}$$

$$G(R_f|R_i; t) = e^{-\tau[V(R_f)-E_T]/2} \, \langle R_f| \, e^{-\tau\hat{T}} \, |R_i\rangle \, e^{-\tau[V(R_i)-E_T]/2} \tag{1.39}$$

$$\widetilde{G}(R_f|R_i; t) = \Psi(R_f)G(R_f|R_i; t)\frac{1}{\Psi(R_i)} \tag{1.40}$$

$$f(R_f) = \int dR_i \widetilde{G}(R_f|R_i; t)\Psi(R_i)^2 \tag{1.41}$$

The stochastic implementation of equation 1.38 can take on different forms but the most popular is the 'birth/death algorithm' under the 'fixed-node approximation'.[159]

Fixed Node Approximation

Fixed Node DMC is a way of dealing with the fermionic sign issue of DMC. [9, 130, 159] The sign issue with the DMC algorithm described above is that it does not take into account the anti-symmetry of fermionic wave functions that we are interested it.[9, 130, 159] Without a constraint to maintain anti-symmetry, the DMC algorithm will project out the bosonic ground state wave function. Even if bosonic states were prohibited both Ψ_0 and $-\Psi_0$ are equally good solutions to the Schrödinger equation and an equal contribution of both from the projection will make the energy evaluate to zero with noise.[9, 130, 159, 14, 206, 63, 79] A simple solution then is to impose that the nodal surface of the trial wave function fix the nodal surface of the projected wave function. [9, 130, 159, 14, 206, 63, 79] In practice this is done by simply having having walkers restricted to the 'nodal pockets' of the wave function.

1.6 Pseudopotentials

Pseudopotentials (PP) are ubiquitous in electronic structure calculations because they improve computational efficiency while producing results similar to all-electron calculations. [77] If errors occur with PP use it is typically expected to be that PP results are not consistent with all electron results.[77, 213] But there are more subtle and lethal errors that can occur when using these PP for QMC.[30, 51] These errors are especially difficult to find the source of due to the stochastic nature of QMC often 'washing them away'. In addition the PP error can be experienced on the QMC side while coming from one or more of the following: 1) PP design, 2) SPO set produced by DFT, 3) how PP is evaluated in QMC.

Throughout the work in this dissertation severe issues arose in the choice and parameterization of PP for calculations that impacted our ability to determine band gaps. It was only

through a careful review of PP that ultimately these issues were resolved to be 'ghost states'. In order to prepare the reader for the QMC specific discussion of this error we give a brief review of the Kleinman-Bylander transformation (which allow for ghost states to occur), and how ghost states occur in DFT to ultimately affect QMC calculations.

Overview

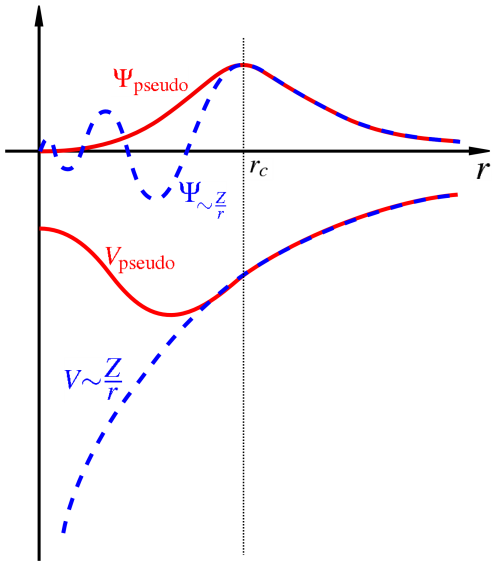


Figure 1.6: Sketch of a 'pseudo wave function' and 'pseudopotential' along with the 'full wave function' and 'full potential'. The $r < r_c$ region refers to the core region [219]

The electrons of an atom can conceptually be divided into the categories of "core" and "valence" in which the noble gas configuration and filled d/f subshells of an electron configuration are typically categorized as core electrons (e.g. for Be the electron configuration is $Be = 1s^2 2s^2$ or $[He], 2s^2$ in which the $[He]$ is the core). It is often stated that only the valence electrons participate in bonding due to the observation that wave functions of valence electrons can change singficantly for a chemical bond while wave functions of core electrons often only change slightly after bond formation. In fact, this simple but powerful concept's influence can be seen in the abstraction of chemical reactions in terms of simple Lewis dot structures employed by organic chemist to teach reactions and build chemical intuition. With respect to electronic structure programs the approximation that core electrons can be ignored when trying to calculate molecular or solid state properties arises in the form of pseudopotentials. These pseudopotentials are useful when dealing with the computational challenge that a large number of filled shells pose in plane wave electronic structure theory.

The idea of a pseudopotential is that the influence of the core electrons and the nucleus of an atom on the valence states can be approximated with an effective potential. The advantages of this approximation are: a reduction of basis size necessary to describe the system, a reduction of the number of electrons present in your calculation, and the ability to simply incorporate some relativistic effects.

A simple sketch of how the pseudopotential produces a nodeless simplified valence wave function is illustrated in figure 1.6. The blue dashed lines represent the all electron single partial orbital valence wave function and full potential, while the red lines represents the pseudo wave function associated with the pseudo potential. The original blue wave function necessarily exhibits high oscillatory character in the core region (defined as $r < r_c$) due to the orthogonality constraint with respect to the core single particle orbital wave functions. This high oscillatory character is what necessitates the use of a large number of fourier modes leading to calculations that are costly for all electron calculations. When the core region of the wave function is simplified to contain no nodes we can imagine that the number of fourier modes needed will be drastically reduced.

In the rest of this section we review the Kleinman-Bylander transformation, ghost states, and finally the use of pseudopotentials in Variational Monte Carlo (VMC).

Semilocal Pseudopotnetial and the Kleinman-Bylander Transformation

The Kleinman-bylander transformation[107] is the logical starting point for our discussion because it is used to evaluate pseudopotentials commonly by modern DFT software. [70]

It is customary to break up the total pseudopotential ($V_{total}(r) = V_{local}(r) + \delta V_l(r)$) into the Semilocal (SL) form of equation 1.42, in which the radial and and angular parts are separated due to spherical symmetry, and the non-locality refers to the angular coordinates. The choice of V_{local} is arbitrary and the sum over l is often truncated at small values.

$$V_{SL} = V_{local}(r) + \sum_{lm} |Y_{lm}\rangle \, \delta V_l(r) \, \langle Y_{lm}| \tag{1.42}$$

The evaluation of the expectation value of equation 1.42 leads to the equation 1.45.

$$e^{iq\cdot r} = 4\pi \sum_{lm} i^l j_l(qr) Y^*_{lm}(q) Y^*_{lm}(r) \tag{1.43}$$

$$\langle q| \, V_{SL} \, |q'\rangle = \frac{(4\pi)^2}{\Omega} \sum_{lm} \sum_{ab} \sum_{cd} \int (-i)^a i^c j_a(qr) j_c(q'r) Y_{ab}(q) Y^*_{cd}(q') V_{nl}(r) r^2 dr \delta_{la,mb} \delta_{lc,md} \tag{1.44}$$

$$\langle q| \, V_{SL} \, |q'\rangle = \frac{4\pi}{\Omega} \sum_{l} (2l+1) \int j_l(qr) j_l(q'r) P_l(cos\theta) V_l(r) r^2 dr \tag{1.45}$$

If the number of plane waves in our basis is N_{pw} this general form leads to the need to calculate N_{pw}^2 integrals.

The Kleinman-Bylander (KB) transformation[107] reduces the number of necessary integrals to just N_{pw} by transforming the semilocal form into a fully non-local (NL) form that is a separable pseudopotential operator.[106]

$$V_{NL} = V_{local}(r) + \sum_{lm} \frac{\left|\delta V_l \psi_{lm}^{PS}\right\rangle \left\langle \psi_{lm}^{PS} \delta V_l\right|}{\left\langle \psi_{lm}^{PS}\right| \delta V_l \left|\psi_{lm}^{PS}\right\rangle} \tag{1.46}$$

$$V_{NL} = V_{local}(r) + \sum_{lm} \left|\eta_l(r)\right\rangle E_l^{KB} \left\langle \eta_l(r)\right| \tag{1.47}$$

The energies E_l^{KB} determine the strength of the nonlocality and can be considered to be coupling constants.

$$E_l^{KB} = \frac{\left\langle \psi_{lm}^{PS}\right| \delta V_l^2 \left|\psi_{lm}^{PS}\right\rangle}{\left\langle \psi_{lm}^{PS}\right| \delta V_l \left|\psi_{lm}^{PS}\right\rangle} \tag{1.48}$$

$$\left|\eta_l(r)\right\rangle = \frac{\left|\delta V_l \psi_{lm}^{PS}\right\rangle}{\sqrt{\left\langle \psi_{lm}^{PS}\right| \delta V_l^2 \left|\psi_{lm}^{PS}\right\rangle}} \tag{1.49}$$

Using V_{NL} on the reference state ψ_{lm}^{PS} is equivalent to operating with the V_{SL}. And we see that only products of projection operations are needed to calculate the expectation value of the potential in our basis leading to our reduction of N_{pw}^2 to N_{pw}.

$$\left\langle \delta V_l \psi_{lm}^{PS} \middle| \psi \right\rangle = \int dr \delta V_l(r) \psi_{lm}^{PS}(r) \psi(r) \tag{1.50}$$

$$\left\langle \psi_i\right| V_{NL} \left|\psi_j\right\rangle = \sum_{lm} \left\langle \psi_i \middle| \psi_{lm}^{PS} \delta V_l \right\rangle \frac{1}{\left\langle \psi_{lm}^{PS}\right| \delta V_l \left|\psi_{lm}^{PS}\right\rangle} \left\langle \delta V_l \psi_{lm}^{PS} \middle| \psi_j \right\rangle \tag{1.51}$$

Ghost states

We are now ready to describe 'ghost states' which produce bad symptoms in a QMC calculation but not always in a DFT calculation. For now we keep the discussion of ghost states limited to DFT calculations but will discuss their consequence in QMC in the next section.

Overview of Ghost states

Although the KB NL PP operator produces the same result as the SL PP operator when applied to the reference state, when the NL PP operator is applied to other states the operation result may change leading to what are known as ghost states.[119] Pseudopotentials that produce eigenvalues of the nonlocal hamiltoninan below the reference energy of the pseudopotential are referred to as ghost states.[73] [103] More simply ghost states are states

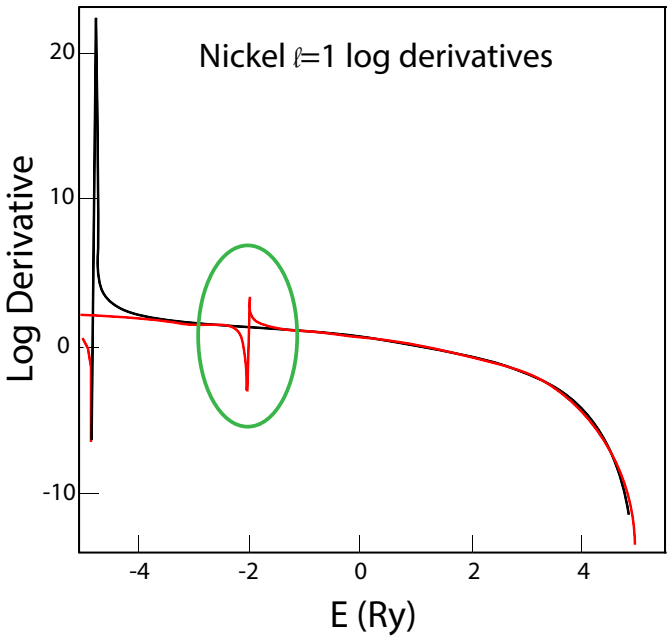


Figure 1.7: Sketch of an example ghost state in the production of pseudopotential for Ni. We see a divergence of $D_l(\epsilon)$ indicating a ghost state.[2]

that violate the Wronskian thoerem [124] (which states that atomic eigenfunctions are energetically ordered such that an increase in the number of nodes corresponds to an increase in the energy). So when ghost states are found it often occurs below the zero-node state of the atom, but they can occur in higher states as well. We will review the theorem that Gonze et al. devised to tell if a a ghost state occurs. [73] [103] One way to detect ghost states is to look at the energy derivative of $D_l(\epsilon)$ and look for large divergences in phase shifts.[103] An example can be seen in figure 1.7 for a Ni atom $l = 1$ pseudowavefunction. But visual inspection of the $D_l(\epsilon)$ will not always reveal the ghost states because the divergences can be quite sharp and missed if small step sizes are not taken.[103]

We will summarize the analysis of separable nonlocal pseudopotentials devised by Alexander Khein in which a method to determine ghost state is described.[103] Beginning with the DFT radial equation the Hamiltonian is broken down to a local component and nonlocal potential component.

$$\begin{aligned}[-\frac{d^2}{dr^2} + \frac{l(l+1)}{r^2} + V_H(r) + V_{xc}(r) + V_{loc}^{ion}(r) \\ + \delta V_l(r)]\psi_{lm}^{PS}(r) = \epsilon_l^{ref}\psi_{lm}^{PS}(r)\end{aligned} \tag{1.52}$$

$$[H_{loc} + \delta V_l(r)]\psi_{lm}^{PS}(r) = \epsilon_l^{ref}\psi_{lm}^{PS}(r) \tag{1.53}$$

$$H = H_{local}(r) + |\eta_l(r)\rangle\, E_l^{KB}\, \langle\eta_l(r)| \tag{1.54}$$

It is understood that there is a radial hamiltonian (equation 1.54) of this form for each angular momentum channel. Now we will work towards deriving the "eigenvalue condition" used to determine if ghost states are present. Although we have shown the projection per channel to be a single operator, it is possible to have more than one projection per angular momentum channel [103][18]. But for the following discussion we will continue to assume a single projection.

Consider the eigenstates of the full Hamiltonian expanded in the basis of the eigenstates of the local Hamiltonian equation 1.55.

$$|\psi\rangle = \sum_m c_m \left|\psi_m^{loc}\right\rangle \tag{1.55}$$

Then we use equation 1.55 and plug it into the Schrödinger equation $H|\psi\rangle = \epsilon|\psi\rangle$

$$E_l^{KB}\sum c_m |\eta_l(r)\rangle \left\langle\eta_l(r)\middle|\psi_m^{loc}\right\rangle = \sum c_m(\epsilon - \epsilon_m^{loc})\left|\psi_m^{loc}\right\rangle \tag{1.56}$$

Taking the inner product with $\left|\psi_k^{loc}\right\rangle$ results in ...

$$c_k = \frac{E_l^{KB}}{\epsilon - \epsilon_m^{loc}}\sum c_m \left\langle\psi_k^{loc}\middle|\eta_l(r)\right\rangle \left\langle\eta_l(r)\middle|\psi_m^{loc}\right\rangle \tag{1.57}$$

Then multiply both sides by $\left\langle\eta\middle|\psi_k^{loc}\right\rangle$, sum over k, and using the fact that

$$\sum_k\ c_k \left\langle\eta\right|\ \left.\psi_k^{loc}\right\rangle = \left\langle \eta\middle|\sum_k\ c_k\ \psi_k^{loc}\right\rangle = \langle\eta|\psi\rangle \tag{1.58}$$

results in equation 1.59, from which we can define the local Hamiltonian's Green's function (eqn. 1.60)

$$\langle\eta_l(r)|\psi\rangle = E_l^{KB}\,\langle\eta_l(r)|\,[\sum_k \frac{\left|\psi_k^{loc}\right\rangle\left\langle\psi_k^{loc}\right|}{\epsilon - \epsilon_k^{loc}}]\,|\eta_l(r)\rangle\,\langle\eta_l(r)|\psi\rangle \tag{1.59}$$

$$\hat{g}(\epsilon) = [\sum_k \frac{\left|\psi_k^{loc}\right\rangle\left\langle\psi_k^{loc}\right|}{\epsilon - \epsilon_k^{loc}}] \tag{1.60}$$

Finally we can rearrange equation 1.59 to be equation 1.61

$$\frac{1}{E_l^{KB}} = \langle\eta_l(r)|\,\hat{g}(\epsilon)\,|\eta_l(r)\rangle \tag{1.61}$$

The $\langle \eta_l(r)| \hat{g}(\epsilon) |\eta_l(r)\rangle$ is plotted in figure 1.8 for some positive value of E_l^{KB}. We see that the local eigenvalues, and non-local eigenvalues can be identified in the graph through the poles of the function and the intersection of $\langle \eta_l(r)| \hat{g}(\epsilon) |\eta_l(r)\rangle$ function with $1/E_l^{KB}$. One of the non-local eigenvalues is necessarily the reference energy of the pseudopotential by construction. For values of ϵ less than ϵ_0^{local} we can see that $\langle \eta_l(r)| \hat{g}(\epsilon) |\eta_l(r)\rangle$ is negative infinity and approaches zero asymptotically from the right.

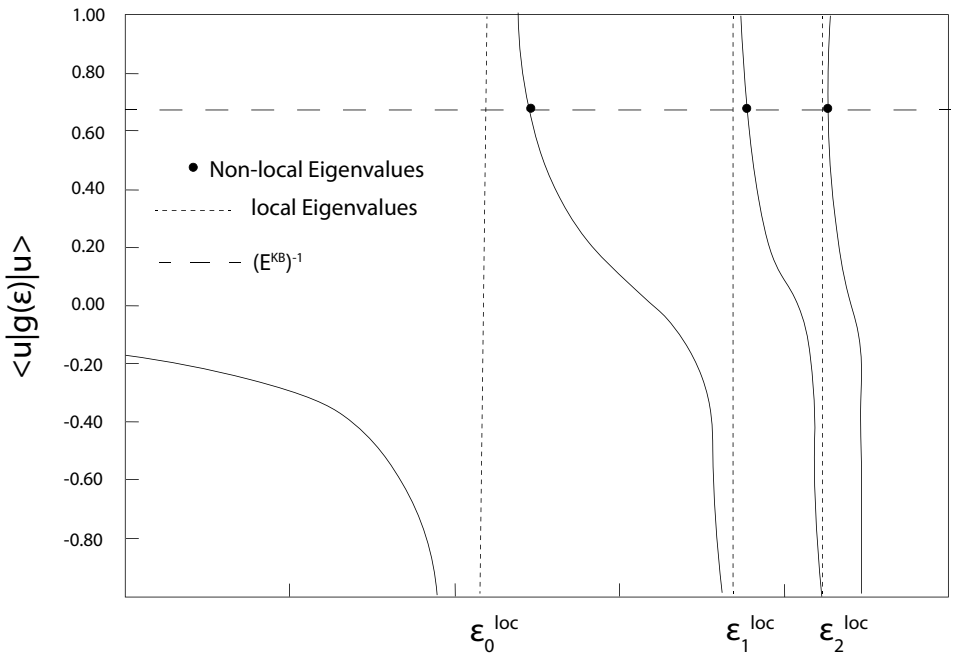


Figure 1.8: $\langle \eta_l(r)| \hat{g}(\epsilon) |\eta_l(r)\rangle$ plotted as a function of ϵ. The value of of E_l^{KB} is positive in this case and the black circles correspond to eigenvalues of the total non-local Hamiltoninan. [103]

We can now summarize the Gonze theorem for detecting ghost states. [103]

1. $E_l^{KB} < 0$: there is a ghost below the reference level E_l if and only if E_l is higher than ϵ_0^{loc}
2. $E_l^{KB} > 0$: there is a ghost below the reference level E_l if and only if E_l is higher than ϵ_1^{loc}

The sign and magnitude of E_l^{KB} seems to be responsible for ghost states. For large negative values of E_l^{KB} we see that a deep lying ghost state can occur. What can be done if a ghost

state is detected? Hopefully some slight modifications to the pseudopotential can resolve the issue. Suggestions include varying the cutoff radius slightly, and changing the local channel potential. These both may alleviate the denominator of equation 1.48 from being too close to zero which is responsible for producing a large E_l^{KB}.

Pseudopotential for Quantum Monte Carlo

Our review of pseudopotentials so far has been exclusively in the independent electron framework, but now we begin discussing the utility and challenges of pseudopotentials in Quantum Monte Carlo (QMC), specifically Variatioanl Monte Carlo (VMC). This section will be divided into three subsections. We first review some basic motivation of PP in QMC and some historical implementations in subsection 1.6. Then we describe some technical details associated with the modern evaluation of the pseudopotential in QMC codes (such as QMCPACK) in subsection 1.6. Finally we touch upon some challenges of pseudopotential design for QMC, and various errors that could arise in calculations in subsection 1.6.

We begin by reminding the reader that VMC involves a trial wave function consisting of a Slater/Multi-Slater portion multiplied by a Jastrow component (a function which explicitly introduces correlation effects and can take many forms). VMC then estimates the ground state energy via a statistical average over N_s position samples $\{\vec{r}_i\}$ drawn from the many-electron probability distribution $p(\vec{r}) = |\Psi(\vec{r})|^2/\langle\Psi|\Psi\rangle$,

$$E = \frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \int p(\vec{r})E_L(\vec{r})d\vec{r} \approx \frac{1}{N_s}\sum_{i=1}^{N_s} E_L(\vec{r}_i), \tag{1.62}$$

$$E_L(\vec{r}) = \frac{H\Psi(\vec{r})}{\Psi(\vec{r})}, \tag{1.63}$$

For more details refer to section 1.5 or QMC review articles. [63] [206]

The need for pseudopotential in QMC

The Variational and Fixed-Node Diffusiom QMC methods are wave functions methods that directly treat quantum many-body effects and involve stochastic sampling for evaluation of quantities. All-electron calculations involving atoms with high Z values presents two problems in getting good statistics for integral evaluations. The first is the that due to the high degree of variation of the wave function at small length scales near the nucleus of atoms, a smaller time step needs to be used for the sampling process, reducing sampling efficiency.[63] The second is that there are large fluctuations of the local energy near the nuclei due to the large values of the kinetic and potential energies of core electrons.[63] Both these issues results in a large sampling effort being necessary to achieve good statistics to evaluate the energy, or perform optimization of the trial wave function. As has been mentioned throughout this review, the core electrons are not necessary to describe inter

atomic bonding and low lying excited states and therefore pseudopotentials could be used to improve computational cost.

In our discussion of the pseudopotentials so far we have shown that generally PP can be written in terms of a local and nonlocal component. If we were to express the pseudopontential in this general form but emphasize the action on each electron i of the total wave function the result would be equation 1.64. [63]

$$V_{loc}(R) + \hat{V}_{nl} = \sum_i V_{loc}^{ps}(r_i) + \sum_i \hat{V}_{nl,i}^{ps} \tag{1.64}$$

When the non-local component of equation 1.64 is applied to an arbitrary function $f(r_i)$ the result is equation 1.65. [63]

$$\hat{V}_{nl,i}^{ps} f(r_i) = \sum_{l,m} V_{nl,l}^{ps}(r_i) Y_{lm}(\Omega_i) \int_{4\pi} Y_{lm}^*(\Omega_i') f(r_i') d\Omega_i' \tag{1.65}$$

The angular momentum components of the function $f(r_i)$ is plucked out by the integrals allowing for each momentum channel to "see" its appropriate potential. [63]

Practical Implementation of pseudopotentials in Variational Monte Carlo

Early implementations of PP in QMC involve just applying the PP to the single slater component of the trail wave function. [78] Fhay, Wang and Louie developed a method to apply the non-local pseudopoential to the full trial wave function (including the correlating Jastrow factor). [54] [55] [63] Remember that the evaluation of the total electronic energy depends on sampling of the local energy ($E_L = \Psi_T^{-1} \hat{H} \Psi_T$) and if we consider a single atom centered at the origin the non-local component of the Hamiltonian is expressed as equation 1.66 in which the sum over i refers to the sum over electrons. [63]

$$\begin{aligned} V_{nl} &= \Psi_T^{-1} \hat{V}_{nl} \Psi_T \\ &= \sum_i \Psi_T^{-1} \hat{V}_{nl,i}^{ps} \Psi_T = \sum_i V_{nl,i} \end{aligned} \tag{1.66}$$

The evaluation of equation 1.66 using equation 1.65 results in equation 1.67.

$$\begin{aligned} V_{nl,i} = \sum_l V_{nl,l}^{ps}(r_i) \sum_{m=-l}^{l} Y_{lm}(\Omega_{r_i}) \int Y_{lm}^*(\Omega_{r_i'}) \\ \times \frac{\Psi_T(r1, ..., r_i', ..., r_N)}{\Psi_T(r1, ..., r_i, ..., r_N)} d\Omega_{r_i'} \end{aligned} \tag{1.67}$$

The angular integration is over a sphere passing through the i'th electron and centered at the origin and can be further simplified to equation 1.68 with a choice of the z-axis being

parallel to r_i (vector from the origin of the atom to the ith electron).[63]

$$V_{nl,i} = \sum_l V^{ps}_{nl,l}(r_i)\frac{2l+1}{4\pi}\int P_l[cos(\theta'_i)] \times \frac{\Psi_T(r1,...,r'_i,...,r_N)}{\Psi_T(r1,...,r_i,...,r_N)}d\Omega_{r'_i} \tag{1.68}$$

Where P_l is the Legendre polynomial and to integrate the products of spherical harmonics in equation 1.68 a quadrature rule is used.

Challenges and errors of pseudopotentials design for QMC

Using QMC to directly design pseduopotentials for QMC calculations has been done and has been shown to be very accurate [4]. But more commonly PPs that were originally designed for single-particle methods are used in QMC. The reason why QMC is not commonly used to directly make PPs is that performing all-electron calculations for an atom with high Z value is costly to achieve sufficient statistical accuracy. But using the independent electron methods such as Hartree Fock or DFT to design a PP (as described in previous sections) to use for QMC introduces errors. Furthermore, it's not even clear if a hierarchy exists for which independent electron method designed PP is most appropriate for use with QMC. It has been shown that often Hartree Fock designed PP yield greater accuracy than DFT PP despite the fact that the DFT method attempts to include correlation effects. Another option to produce PP for QMC is to use quasiparticle / many-body methods (such as Couple Cluster Singles and Doubles, and GW) to produce PP.[17] [194]

For the independent electron framework we have previously discussed, the separation of core and valence electrons is straightforwardly done by partitioning the orbitals as either associated with the core electron or valence electrons. [63] But within the many electron framework (such as QMC) the core-valence partition is not exact because electrons are indistinguishable particles. Furthermore a "removal of the core" removes correlation energy between core electrons, and correlation energy between core and valence electrons. [63] These errors can be large if the number of valence electrons is small relative to the core because changes in the small valence can cause core polarization and relaxations. 'Core polarization potentials' (CPP) have been used to modify the Hamiltonian and to address these issues [63] [112] [195] and can be determined both from first principles and using simple empirical forms. [63] [133] It has been shown that the CPP significantly improve QMC accuracy. [196]

The choice of l_{max} involved in the sum of angular momentum components in equation 1.65 is another source of possible errors when employing pseudopotentials for QMC. Typically in the independent electron framework the choice of l_{max} is based off the highest orbital occupancy of electrons in an atom.[63] For example, the oxygen atom has electrons occupy s and p orbitals therefore $l_{max} = 1$. We shall refer to this rationale for choice of l_{max} as the 'standard method'. But it must be considered that the Jastrow factor can introduce higher angular momentum components that the orbitals in the determinatal part of the wave

function does not contain.[63] An extensive study of the importantance of higher angular momentum channels for PP in QMC has been done by Tipton et. al. [205] They found that for QMC the choice of l_{max} and what angular momentum is chosen for the local channel both have a strong influence on the accuracy of the calculation. One observation made is that if the pseudopoential contains angular momentum terms higher than that determine via the 'standard method' resulting VMC energies are less sensitive to the choice for the local channel. But when l_{max} was determined via the 'standard method', different choices for the local channel led to VMC energy differences that could range from a few to tens of milliHartree. In addition, the VMC variances of the calculation could also vary to as much as 0.3 Hartree squared. This "variance" dependence on the local channel potential can have a strong bearing on results of "variance matched" calculations that depend on matching the variance between two different states to have a greater cancellation of error when determining quantities such as vertical excitation energies or reaction barriers. [160] [155] (We discuss our variance matching scheme in more detail in section 2.4.) The explanation for this sensitivity to the local channel choice is that the higher momentum components of the wave function are present but do not experience their appropriate l-dependent potential, instead only experiencing the local channel potential. When the local channel happens to be chosen to be the high angular momentum, errors are mitigated. But when the local channel is chosen to have a $l = 0$ value then the high angular momentum components of the wave function experience an incorrect potential.

Although ghost states are specific to using the completely nonlocal Kleinman-Bylander representation of pseudopotentials in plane wave DFT codes, QMC uses the single particle orbitals from these DFT calculations to construct a trial wave function and is therefore sensitive to ghost state issues as well. Drummond, Trail and Needs performed an analysis of of Dirac-Fock pseudopotentials for QMC and found serious problems due to ghost states are quite common when using DFT generated orbitals for VMC.[50] As stated in the article the consequences of ghost states for a plane-wave-DFT-QMC workflow can be the following...[50]

> The presence of ghost states gives rise to some or all of the following symptoms: the failure of the DFT self-consistentfield (SCF) process to converge; a large difference between the DFT energies obtained with plane-wave and Gaussian basis sets; the existence of an absurdly low Kohn-Sham eigenvalue; an absurdly high (unbound) energy when the orbitals are used in VMC calculations; a very large energy variance; enormous difficulty optimizing a trial wave function in VMC; and enormous difficulty controlling the configuration population in a DMC simulation. Furthermore, these difficulties may change or disappear when the local channel is changed.

We now summarize other findings from the article and note that explanations for these observations are not offered by the author. The article describes how it is also possible for a DFT calculation to not seem to suffer from a ghost state issue while a QMC calculation exhibits symptoms mentioned. There was also an observation that high plane wave cut off

energies are needed sometimes otherwise VMC optimization issues and variance values tend to be large. They tested the cut off energy necessary to achieve chemical accuracy (1.59 mHa) and an order of magnitude greater accuracy (0.1 mHa) for convergence of the DFT total energy of atoms. They found that for transition metals the cut off energy could be almost as high as 10^5 Ha which suggest that caution must be taken if transition metals are present in our plane-wave-DFT-QMC workflow. Another observation of the article is that to mitigate the presence of ghost state one should use and s-local channel for DFT calculations while using a d-local channel for QMC calculations and this is what we have done for our calculations.

Chapter 2

QMC Specific Theory and Methods

2.1 Excited State Target Function

To ensure our orbitals are tailored to the needs of an individual excited state, we will rely on the excited state variational principle that minimizes the objective function Ω,

$$\Omega = \frac{\langle\Psi|(\omega - H)|\Psi\rangle}{\langle\Psi|(\omega - H)^2|\Psi\rangle} \tag{2.1}$$

which is a function whose global minimum is the exact Hamiltonian eigenstate with energy immediately above the value ω. [227] Just as ground state VMC estimates the energy via a statistical average over N_s position samples $\{\vec{r}_i\}$ drawn from the many-electron probability distribution $p(\vec{r}) = |\Psi(\vec{r})|^2/\langle\Psi|\Psi\rangle$,

$$E = \frac{\langle\Psi|H|\Psi\rangle}{\langle\Psi|\Psi\rangle} = \int p(\vec{r})E_L(\vec{r})d\vec{r} \approx \frac{1}{N_s}\sum_{i=1}^{N_s} E_L(\vec{r}_i), \tag{2.2}$$

$$E_L(\vec{r}) = \frac{H\Psi(\vec{r})}{\Psi(\vec{r})}, \tag{2.3}$$

the objective function Ω may be statistically estimated as a ratio of two such averages

$$\Omega \approx \frac{\sum_{i=1}^{N_s} \omega - E_L(\vec{r}_i)}{\sum_{i=1}^{N_s}(\omega - E_L(\vec{r}_i))^2} \tag{2.4}$$

and minimized via generalizations [227, 226, 190] of the ground state Linear Method. [215, 207] Note that, in practice, it is advisable to use a slightly modified probability distribution from which to draw the samples, a point we will return to in Section 2.2.

While this and other [125, 41] excited state variational principles are quite general, one of their most promising uses is to help achieve excited-state-specific relaxations of the orbital basis. While excited state variational principles have been employed deterministically in both

single-determinant wave functions [222] and linear combinations of single excitations, [191] as well as by VMC for Jastrow-modified single excitations, [19, 20] the prospect of extending the approach to the class of highly-sophisticated multi-Slater Jastrow (MSJ) wave functions that are directly compatible with VMC and DMC is quite intriguing. In the following sections, we will discuss how these excited state variational principles may be combined with other recent advances in pursuit of this goal.

2.2 Modified Guiding Function

Although ground state VMC often draws samples from the probability distribution $|\Psi|^2/\langle\Psi|\Psi\rangle$ due to the allure of the zero variance principle, [64] this approach is not statistically robust when estimating the energy variance,

$$\sigma^2 = \frac{\langle\Psi|(H-E)^2|\Psi\rangle}{\langle\Psi|\Psi\rangle}. \tag{2.5}$$

The trouble comes from the fact that the local energy $H\Psi/\Psi$ can diverge, because Ψ can be zero when $\nabla^2\Psi$ is not. Although this divergence is integrable for E and σ^2 and so poses no formal issues for estimating E, it is not integrable for the variance of σ^2, [212, 211, 161] and so a naive approach in which samples are drawn from $|\Psi(\vec{r})|^2/\langle\Psi|\Psi\rangle$ will not produce normally distributed estimates for σ^2. Due to the relationship

$$\Omega(\Psi) = \frac{\langle\Psi|(\omega-\hat{H})|\Psi\rangle}{\langle\Psi|(\omega-\hat{H})^2|\Psi\rangle} = \frac{\omega-E}{(\omega-E)^2+\sigma^2} \tag{2.6}$$

we are left with the consequence that statistical estimates for Ω via Eq. (2.7) will also not be normally distributed.

$$\Omega \approx \frac{\sum_{i=1}^{N_s} \omega - E_L(\vec{r}_i)}{\sum_{i=1}^{N_s}(\omega - E_L(\vec{r}_i))^2} \tag{2.7}$$

In a previous study, [161] we overcame this difficulty with the alternative importance sampling function

$$|\Psi_{\mathrm{M}}|^2 = |\Psi|^2 + \frac{\epsilon\,|\nabla^2\Psi|^2}{1+\exp\left[\left(\ln|\Psi| - \overline{\ln|\Psi|} + \sigma_{\overline{\Psi}}\right)/\sigma_{\overline{\Psi}}\right]} \tag{2.8}$$

in which the average ($\overline{\ln|\Psi|}$) and standard deviation ($\sigma_{\overline{\Psi}}$) of the logarithm of the wave function absolute value are estimated on a short sample drawn from the traditional distribution $|\Psi(\vec{r})|^2/\langle\Psi|\Psi\rangle$. For any $\epsilon > 0$, Eq. (2.8) guarantees that the wave function ratio

$$\eta(\vec{r}) = \left|\Psi(\vec{r})/\Psi_{\mathrm{M}}(\vec{r})\right| \tag{2.9}$$

and the modified local energy

$$E_L^{\mathrm{M}}(\vec{r}) = \eta(\vec{r}) E_L(\vec{r}) \tag{2.10}$$

will be finite everywhere, which implies that if we draw N_s samples from the distribution $|\Psi_{\mathrm{M}}(\vec{r})|^2/\langle\Psi_{\mathrm{M}}|\Psi_{\mathrm{M}}\rangle$, the resulting statistical estimates

$$E \approx \frac{\sum_{i=1}^{N_s} \eta(\vec{r}_i) E_L^{\mathrm{M}}(\vec{r}_i)}{\sum_{i=1}^{N_s} \left(\eta(\vec{r}_i)\right)^2} \tag{2.11}$$

$$\sigma^2 \approx \frac{\sum_{i=1}^{N_s} \left(E_L^{\mathrm{M}}(\vec{r}_i) - \eta(\vec{r}_i) E\right)^2}{\sum_{i=1}^{N_s} \left(\eta(\vec{r}_i)\right)^2} \tag{2.12}$$

$$\Omega \approx \frac{\sum_{i=1}^{N_s} \eta(\vec{r}_i)\Big(\omega\,\eta(\vec{r}_i) - E_L^{\mathrm{M}}(\vec{r}_i)\Big)}{\sum_{i=1}^{N_s} \Big(\omega\,\eta(\vec{r}_i) - E_L^{\mathrm{M}}(\vec{r}_i)\Big)^2} \tag{2.13}$$

are guaranteed to be normally distributed for sufficiently large N_s. Note that, due to divergences in E_L, this central limit theorem guarantee would not be true [212, 211] for the σ^2 and Ω estimates if we had made the traditional choice of $|\Psi_{\mathrm{M}}|^2 = |\Psi|^2$. Also note that, although the denominator in Eq. (2.8) is not strictly necessary in order to recover normal statistics for the estimates of σ^2 and Ω, it does help keep us as close as possible to $|\Psi|^2$ and thus the zero variance principle by smoothly switching off the modification when the value of the wave function magnitude is large relative to its average. This way, the divergences that occur near the nodes of Ψ are avoided, while at the same time the probability distribution is left essentially unmodified in regions of space where the wave function magnitude is large.

By exploiting the table method [42], it is possible to employ Ψ_M without changing the overall cost scaling of the Markov chain propagation. Although we now must evaluate $\nabla^2\Psi = -2\Psi K_L$ every time we move one of our electrons, the local kinetic energy K_L has the same cost scaling as that of evaluating Ψ itself once the matrices $\boldsymbol{A}$, $\boldsymbol{B}$, $\boldsymbol{T}$, and $\boldsymbol{M}$ (see section 2.5) have been prepared. Thanks to the Sherman Morrison formula, these matrices can be updated efficiently during each one-electron move, and although the new per-move need for $\boldsymbol{M}$ and K_L does increase the update cost, it does not change the scaling. Overall, our experience has been that the practical benefits of using Ψ_{M} to achieve normally distributed estimates for σ^2 and Ω more than make up for the additional cost of its Markov chain propagation.

Before discussing our method's efficacy in predicting excitation energies, we would like to emphasize the benefit of drawing samples from the modified guiding wave function Ψ_{M}. As seen in Figure 2.1, even the relatively simple optimization of a 6-configuration wave function for the first excited singlet of thioformaldhyde benefits significantly from the recovery of normal statistics for our estimates of Ω. For a sample size of $N_s = 768,000$ drawn from either $|\Psi_{\mathrm{M}}|^2$ or $|\Psi|^2$, the worst uncertainties seen in Ω during the last 25 iterations (as

measured by a blocking analysis that assumes the statistics are normal) are a factor of 4 smaller for the modified guiding case. Even if $|\Psi|^2$ resulted in normal statistics (which it does not), this would imply that our modified guiding function reduces the number of samples needed to reach a given uncertainty by a factor of 16, which more than makes up for the roughly 3 to 4 times increased cost per sample of propagating the Markov chain for $|\Psi_{\mathrm{M}}|^2$. It is also worth noting that, thanks to the decreased uncertainty, the optimization that employed $|\Psi_{\mathrm{M}}|^2$ (which is used as the guiding function not only when estimating Ω but also when evaluating the derivatives needed by the linear method) was able to converge to a lower average value of Ω. For more details on thioformaldhyde, see Section 3.2.

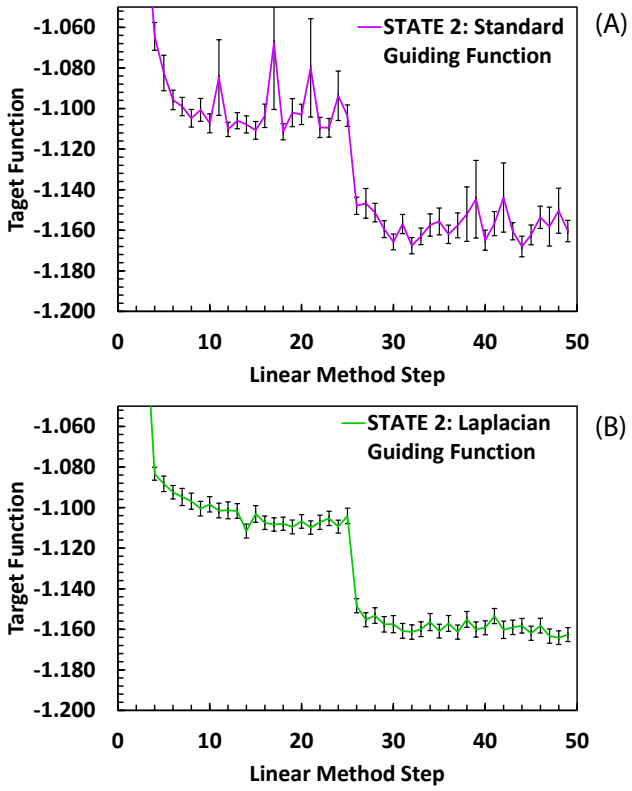


Figure 2.1: The objective function Ω (in $(E_{\mathrm{h}})^{-1}$) during the optimization of the first singlet excited state of SCH_2 using a 6 configuration wave function. Here we compare results using the traditional $|\Psi|^2$ guiding function (A) and our modified guiding function $|\Psi_{\mathrm{M}}|^2$ (B). The first 25 steps hold the orbitals fixed while optimizing the CI and Jastrow parameters while the last 25 steps optimize the orbitals as well.

2.3 Configuration Selection

Two different methods were used to produce configurations for MSJ wave functions. For the aperiodic systems we use S-CI, while for the periodic systems we use a method we dub 'VMC-CI'.

Configuration selection for aperiodic systems

The rapid progress in selective CI methods in recent years has greatly simplified the selection of configurations for MSJ wave functions. Although we expect that any modern selective CI method would work well with our approach, we have taken advantage of existing links between the QMCPACK code [104] and the CIPSI implementation within Quantum Package [1] in order to extract configurations from the CIPSI variational wave function. As studied by Dash et al, [45] a MSJ wave function can either be arrived at by stopping the CIPSI algorithm once its expansion has reached the desired configuration number, or by intentionally running CIPSI to a much larger configuration number and then truncating to the number desired for use in MSJ. Following their recommendation that the latter method is more effective, we have for each of our systems iterated CIPSI with all non-core electrons and orbitals active until each state's variational wave function contains at least 5,000 configurations, after which we truncate to the (typically much smaller) set of configurations used in a state's MSJ wave function by retaining the configurations with the highest CIPSI weights for that state. Although there is no guarantee that all of the first so many configurations from this procedure will be needed after introducing Jastrow factors and orbital relaxations, we expect that this strategy of converging CIPSI to a much larger number of configurations than needed will help identify the most important configurations. Although 5,000 configurations is far too few for even perturbatively-corrected CIPSI to be converged for most of the systems we consider, the subsequent addition of state-specific orbital optimization, Jastrow factors, and variance matching allows this lightweight approach to be quite accurate.

Configuration selection for periodic system

The VMC-CI method refers to 'configuration interaction like' calculation being performed in the presence of a Jastrow factor. The steps are as follows...

1. Optimize the Jastrow Factor associated with a SJ wave function for the ground state.

2. Include all potential configurations one would like to consider in a linear expansion of the wave function. In particular we begin with a CIS expansion, then select for the most heavily weighted configurations as a starting point to produced double excitation configurations. This procedure will be referred to as QMC-CIS(D).

3. Evaluate the Hamiltonian matrix by performing Monte Carlo integration of the matrix elements. Then diagonalize the matrix to have ground state and excited state eigenweights of the configuration

4. With the initial weights of these configurations determined, now optimize all parameter in the wave function variationally via for both the ground and excited state.

2.4 Variance Matching

Two different methods were used for 'variance matching'. For the aperiodic systems we use an 'implicit' method, while for periodic systems the variance of the ground and excited state were explicitly matched.

Variance matching aperiodic systems

In our previous work [161] we showed that, in practice, predictions of energy differences can be improved by adjusting the sizes of different states' MSJ CI expansions such that the states' energy variances σ^2 were equal. The idea is to exploit the fact that σ^2 is essentially a measurement for how close a state is to being a Hamiltonian eigenstate, and, in the absence of a more direct measure of a states' energy error, this measurement should be useful in ensuring that different states are modeled at similar levels of quality so as to avoid bias. That this approach helps improve cancellation of error is likely due, at least in part, to the fact that the energies of low-lying states tend to converge from above for large CI expansions, as the missing tail of small-coefficient determinants means that what tends to be missing is a full accounting of weak correlation effects, which in low-lying states tend to lower a state's energy.

As before, we take the approach of evaluating both E and σ^2 for a series of ground state wave functions of differing CI expansion lengths so that we can interpolate to the expansion length for which the ground state variance matches that of the excited state. We perform the interpolation via the nonlinear fitting function (NLFF)

$$f(N) = c + \frac{d}{N^\alpha} \tag{2.14}$$

where the functional form f is used to interpolate both the energy E and energy variance σ^2 by fitting the values c, d, and α for each case separately based on an uncertainty-weighted least-squares fit. Once these fits are made, we can estimate the expansion length N for which the ground state variance would match that of the excited state, and then, for that value of N, what we expect the ground state energy would be. Note that this is only one possible approach, as we could equally well have fixed the ground state wave function and varied the number of determinants in the excited state. An example of the variance and energy matching fits and procedure (with switched roles for ground and excited states) for

CH_2NH is provided in Section 3.1 Figure 3.1. In some cases (e.g. see Section 3.3) it makes more sense to seek an explicit match between two states' variances rather than relying on interpolation.

Variance matching periodic systems

Implicit variance matching scheme works well for aperiodic systems because the Multi-Slater expansions achievable correspond to significant improvements in the variance, but this is not the case for periodic systems. As we noted in section 1.3 much larger expansions of the wave functions is necessary for periodic systems to achieve the same gains that are seen in aperiodic systems (because of the necessary sampling of the Brillouin zone). To illustrate this point consider an aperiodic system with N electrons with a corresponding peridodic system whose minimal unit cell contained N electrons and requires a 4x4x4 Monkhorst-Pack k-point grid to sufficiently sample the Brillouin zone. Now lets suppose that the aperiodic system necessitates an active space of 4 electrons in 4 orbitals for desired accuracy, the corresponding active space for the periodic system would be 256 electrons in 256 orbitals. This is far beyond what active space methods can currently achieve (see section 1.3).

So as an alternative to implicit variance matching we perform explicit variance matching which is simply having the ground and excited state energy variances match. In practice if the variances of the ground and excited state are within one standard error of each other this is considered variance matched.

2.5 Orbital Optimization of Multi-Slater Jastrow Wave Functions

The Table Method

First introduced by Clark and coworkers [42, 127] and recently improved by Filippi and coworkers, [62, 12] the table method has dramatically increased the size of CI expansions that can be handled by VMC within a MSJ wave function. While the reader is encouraged to consult the above publications for a fully detailed explanation of the table method, we will review the theory here as it will prepare us for the discussion of applying automatic differentiation. To understand the table method's efficacy, we will analyze the MSJ wave function

$$\Psi(\vec{r}) = \Phi(\vec{r})\, \psi_J(\vec{r}) \tag{2.15}$$

$$\Phi(\vec{r}) = \sum_{I=0}^{N} c_I D_I(\vec{r}) \tag{2.16}$$

in which ψ_J is the symmetric Jastrow correlation factor and Φ the linear combination of antisymmetric Slater determinants D_I. While in practice one can (and our software does)

exploit the factorization $D_I = D_{I\uparrow} D_{I\downarrow}$ in situations in which the number of electrons of each spin is fixed, we will for ease of presentation describe the theory in terms of a fictitious system in which all the electrons are spin up, in which case Eq. (2.16) applies without further factorization of D_I. The generalization to cases where electrons of both spins are present is straightforward if a bit tedious.

In our system of n up-spin electrons and m orbitals, we may define an $n \times m$ matrix $\boldsymbol{A}$ whose elements are the orbital values for each electron's position,

$$A_{i,j} = \phi_j(r_i) \tag{2.17}$$

with the first n columns corresponding to the orbitals in the reference determinant D_0. Note that we are using notation where $\vec{r}$ is a length $3n$ vector of all the electron coordinates, while r_i is a length 3 vector of the ith electron's coordinates. We write the determinants in our multi-Slater (MS) expansion as

$$D_I = \det(\boldsymbol{A}_I) \tag{2.18}$$

where the $n \times n$ matrix $\boldsymbol{A}_I$ is formed by taking only those columns of $\boldsymbol{A}$ that correspond to orbitals that are occupied in the Ith electron configuration. We will designate the $I = 0$ configuration as the reference configuration, typically the Aufbau configuration, so that, starting from the matrix $\boldsymbol{A}_0$, we may construct the matrix $\boldsymbol{A}_I$ via k_I column replacements, where k_I is the number of single-electron excitations required to transform configuration 0 into configuration I. Using two $n \times k_I$ matrices $\boldsymbol{U}_I$ and $\boldsymbol{P}_I$, one can express this relationship as

$$\boldsymbol{A}_I = \boldsymbol{A}_0 + \boldsymbol{U}_I \boldsymbol{P}_I^T. \tag{2.19}$$

Specifically, each column of $\boldsymbol{P}_I$ has one element with value 1 and the rest zero, while each column of $\boldsymbol{U}_I$ contains the difference between the column of $\boldsymbol{A}$ needed for the Ith configuration and the one it replaces from the reference configuration. Noting that $\boldsymbol{P}_I^T \boldsymbol{P}_I$ is the $k_I \times k_I$ identity matrix, we can rearrange Eq. (2.19) as

$$\boldsymbol{U}_I = (\boldsymbol{A}_I - \boldsymbol{A}_0)\boldsymbol{P}_I, \tag{2.20}$$

which, along with the matrix determinant lemma, allows us to write D_I in terms of D_0 and the determinant of a $k_I \times k_I$ matrix $\boldsymbol{\alpha}_I$,

$$\boldsymbol{\alpha}_I = \boldsymbol{P}_I^T \boldsymbol{A}_0^{-1} \boldsymbol{A}_I \boldsymbol{P}_I \tag{2.21}$$

$$\begin{aligned} D_I &= \det(\boldsymbol{A}_I) \\ &= \det(\boldsymbol{A}_0) \det(\boldsymbol{I} + \boldsymbol{P}_I^T \boldsymbol{A}_0^{-1} \boldsymbol{U}_I) \\ &= D_0 \det(\boldsymbol{I} + \boldsymbol{P}_I^T \boldsymbol{A}_0^{-1} (\boldsymbol{A}_I - \boldsymbol{A}_0)\boldsymbol{P}_I) \\ &= D_0 \det(\boldsymbol{\alpha}_I). \end{aligned} \tag{2.22}$$

As originally recognized by Clark et al, [42] $\boldsymbol{\alpha}_I$ can be constructed efficiently by simply copying the appropriate elements from the precomputed $n \times m$ "table" matrix

$$\boldsymbol{T} = \boldsymbol{A}_0^{-1}\boldsymbol{A} \tag{2.23}$$

from whence the table method takes its name. Thus the cost of evaluating the contribution each additional configuration makes to the wave function value goes as only $(k_I)^3$, which since in practice k_I tends to be small represents a large speedup compared to the n^3 per-configuration cost that would be incurred if the different determinants D_I were evaluated directly as $\det(\boldsymbol{A}_I)$. All together, the wave function logarithm can now be expressed as

$$\ln(\Phi) = \ln D_0 + \ln Q \tag{2.24}$$

$$Q = c_0 + \sum_{I=1}^{N} c_I \det(\boldsymbol{\alpha}_I) \tag{2.25}$$

where $\ln D_0$ is the result in the single-Slater case and $\ln Q$ corrects for the presence of additional configurations.

Following the presentation of Filippi and coworkers, [62, 12] we can see how this efficiency can be extended to evaluating the local energy by defining a Jastrow-dependent one-body operator $\hat{O}$.

$$\hat{O}_i = -\frac{1}{2}\left(\frac{\nabla_i^2\psi_J}{\psi_J} + \frac{2\nabla_i\psi_J \cdot \nabla_i}{\psi_J} + \nabla_i^2\right) \tag{2.26}$$

$$\hat{O} = \sum_{i=1}^{n} \hat{O}_i \tag{2.27}$$

By forming the intermediates

$$\begin{aligned} t_I &= \sum_{i=1}^{n} -\frac{1}{2}\frac{\nabla_i^2(\psi_J D_I)}{\psi_J D_I} \\ &= \frac{1}{D_I}\sum_{i=1}^{n} -\frac{1}{2}\left[\frac{\nabla_i^2\psi_J}{\psi_J} + \frac{2\nabla_i\psi_J \cdot \nabla_i}{\psi_J} + \nabla_i^2\right] D_I \\ &= \frac{1}{D_I}\sum_{i=1}^{n} \hat{O}_i D_I \\ &= \frac{\hat{O}D_I}{D_I} \end{aligned} \tag{2.28}$$

the kinetic part of the local energy E_L can be written as

$$K_L = \frac{\sum_I c_I D_I t_I}{\sum_I c_I D_I}. \tag{2.29}$$

The kinetic energy intermediates t_I can be converted into a particularly convenient form by defining the $n \times m$ matrix $\boldsymbol{B}$ with elements

$$B_{i,j} = \hat{O}_i \, \phi_j(r_i), \tag{2.30}$$

from which $n \times n$ matrices $\boldsymbol{B}_I$ for each configuration can be constructed in the same fashion as the matrices $\boldsymbol{A}_I$ were derived from $\boldsymbol{A}$. Crucially, one can now use the Leibniz formula to rewrite the intermediates as

$$t_I = \frac{\hat{O} D_I}{D_I} = \frac{\partial}{\partial \lambda} \ln\big(\det(\boldsymbol{A}_I + \lambda \boldsymbol{B}_I)\big)\Big|_{\lambda=0}. \tag{2.31}$$

Now, for a generic invertable matrix $\boldsymbol{G}$ with cofactor matrix $\boldsymbol{C}$, one can use the cofactor formulas for the determinant and the matrix inverse to arrive at the identity

$$\begin{aligned}
\frac{\partial}{\partial \xi} \ln(\det(\boldsymbol{G})) &= \sum_{i,j} \frac{\partial \ln(\det(\boldsymbol{G}))}{\partial G_{i,j}} \frac{\partial G_{i,j}}{\partial \xi} \\
&= \sum_{i,j} \frac{C_{i,j}}{\det(\boldsymbol{G})} \frac{\partial G_{i,j}}{\partial \xi} \\
&= \sum_{i,j} G^{-1}_{j,i} \frac{\partial G_{i,j}}{\partial \xi} \\
&= \mathrm{Tr}[\boldsymbol{G}^{-1} \frac{\partial \boldsymbol{G}}{\partial \xi}].
\end{aligned} \tag{2.32}$$

For the reference configuration, this identity gives us

$$t_0 = \mathrm{Tr}[\boldsymbol{A}_0^{-1} \boldsymbol{B}_0]. \tag{2.33}$$

For the other configurations, we note that Eq. (2.22) remains valid under the replacement $\boldsymbol{A} \to \boldsymbol{A} + \lambda \boldsymbol{B}$, which we use with Eqs. (2.31) and (2.32) to find that

$$\begin{aligned}
t_I &= \frac{\partial}{\partial \lambda} \Big[\ln\big(\det(\boldsymbol{A}_0 + \lambda \boldsymbol{B}_0)\big) \\
&\quad + \ln\Big(\det\big(\boldsymbol{P}_I^T (\boldsymbol{A}_0 + \lambda \boldsymbol{B}_0)^{-1} (\boldsymbol{A}_I + \lambda \boldsymbol{B}_I) \boldsymbol{P}_I\big)\Big) \Big]_{\lambda=0} \\
&= t_0 + \mathrm{Tr}[\boldsymbol{\alpha}_I^{-1} \boldsymbol{\beta}_I]
\end{aligned} \tag{2.34}$$

where we have defined the $k_I \times k_I$ matrix

$$\begin{aligned}
\boldsymbol{\beta}_I &= \frac{\partial}{\partial \lambda} \Big(\boldsymbol{P}_I^T (\boldsymbol{A}_0 + \lambda \boldsymbol{B}_0)^{-1} (\boldsymbol{A}_I + \lambda \boldsymbol{B}_I) \boldsymbol{P}_I \Big)\Big|_{\lambda=0} \\
&= \boldsymbol{P}_I^T (\boldsymbol{A}_0^{-1} \boldsymbol{B}_I - \boldsymbol{A}_0^{-1} \boldsymbol{B}_0 \boldsymbol{A}_0^{-1} \boldsymbol{A}_I) \boldsymbol{P}_I.
\end{aligned} \tag{2.35}$$

As for $\boldsymbol{\alpha}_I$ and $\boldsymbol{T}$, we can define a second table matrix

$$\boldsymbol{M} = \boldsymbol{A}_0^{-1}\boldsymbol{B} - \boldsymbol{A}_0^{-1}\boldsymbol{B}_0\boldsymbol{A}_0^{-1}\boldsymbol{A} \tag{2.36}$$

such that each $\boldsymbol{\beta}_I$ can be built by simply copying the appropriate elements from the pre-computed matrix $\boldsymbol{M}$. Combining Eqs. (2.22), (2.25), (2.29), and (2.34) leads us to our final expression for the kinetic portion of the local energy

$$K_L = t_0 + \frac{1}{Q}\sum_{I=1}^{N} c_I \det(\boldsymbol{\alpha}_I)\,\mathrm{Tr}[\boldsymbol{\alpha}_I^{-1}\boldsymbol{\beta}_I] \tag{2.37}$$

in which the local kinetic energy t_0 of the reference-configuration-based single-Slater-Jastrow wave function is corrected by the second term to produce the local kinetic energy of the full MSJ wave function. We therefore see that the table method allows the local energy to be evaluated for a cost that goes as n^2m for the construction of $\boldsymbol{T}$, $\boldsymbol{M}$, and t_0 plus an additional per-configuration cost that goes as just $(k_I)^3$.

To achieve state-specific orbital optimization, we minimize the excited state objective function Ω via a generalization of the linear method. [190] In practice, this requires evaluating the derivatives of E_L and $\ln(\Psi)$ with respect to the wave function variables at every sample of the electron positions, an endeavor that has recently been made drastically more efficient thanks to the approach of Filippi and coworkers. [62, 12] In their derivation, they construct the matrix

$$\Gamma = \left(\frac{\partial \ln \Phi}{\partial \boldsymbol{A}}\right)^T \tag{2.38}$$

and reformulate the local kinetic energy as

$$K_L = \mathrm{Tr}\left[\Gamma \boldsymbol{B}\right], \tag{2.39}$$

which they then combine with efficient derivatives of Γ to provide a general and efficient approach for determining wave function and local energy derivatives with respect to arbitrary variables, such as orbital parameters or nuclear coordinates. Compared to the original table method, this approach greatly reduces the cost of evaluating these derivatives.

Automatic Differentiation

Inspired by Filippi's efficiency breakthrough, we set out to discover whether there would be any advantage to instead directly applying AD to the wave function and local kinetic energy expressions. In short, we found that the AD route leads to a formulation very similar to Filippi's in which the terms that involve summing over the MSJ's configurational expansion are identical, and so it does not seem to offer any practical advantages. It is however interesting to confirm that AD essentially agrees with the existing method.

To get started with an AD approach, we recognize that, for a given set of electron positions, both K_L and $\ln(\Phi)$ are many-input/single-output functions of the wave function's variational parameters, and so we expect that the AD approach of reverse accumulation [75, 201, 191] will yield all derivatives of these functions for a cost that is a small constant multiple of the cost to evaluate the functions themselves. Reverse accumulation is essentially a careful exploitation of the chain rule, and so, starting with Eqs. (2.24) and (2.25) and recalling that $\boldsymbol{\alpha}_I$ is a $k_I \times k_I$ matrix with elements copied from $\boldsymbol{T}$, we formulate the wave function derivative with respect to a variable μ as

$$\frac{\partial \ln \Phi}{\partial \mu} = \frac{\partial \ln D_0}{\partial \mu} + \mathrm{Tr}\left[\left(\frac{\partial \ln Q}{\partial \boldsymbol{T}}\right)^T \frac{\partial \boldsymbol{T}}{\partial \mu}\right]. \tag{2.40}$$

The first term here is the derivative for a single-Slater wave function, with the second term giving the correction for the multi-Slater case. Once the derivative matrix $\partial \ln Q/\partial \boldsymbol{T}$ has been found, reverse accumulation can continue via the back propagation graph of $\boldsymbol{T}$ itself at a cost that is independent of the number of configurations N. Note especially that the cost and difficulty of this stage of reverse accumulation is not made substantially harder if we wish to have derivatives with respect to multiple types of variables, such as orbital rotations and nuclear positions, as reverse accumulation graphs for simple matrix expressions like Eq. (2.23) and the construction of $\boldsymbol{A}$ from atomic orbitals are straightforward. Crucially, reverse accumulation does not require us to explicitly form $\partial \boldsymbol{T}/\partial \mu$ for all the different quantities μ we wish to differentiate with respect to, so the key challenge in this approach is the evaluation of the derivatives $\partial \ln Q/\partial \boldsymbol{T}$ that are to be fed in to the rest of the back propagation computational graph.

To evaluate these derivatives $\partial \ln Q/\partial \boldsymbol{T}$, we use the determinant derivative

$$\frac{\partial}{\partial x}\det(\boldsymbol{\alpha}_I) = \det(\boldsymbol{\alpha}_I)\mathrm{Tr}\left[\boldsymbol{\alpha}_I^{-1}\frac{\partial \boldsymbol{\alpha}_I}{x}\right] \tag{2.41}$$

with Eq. (2.25) to get

$$\frac{\partial \ln Q}{\partial T_{ij}} = \frac{1}{Q}\sum_{I=1}^{N} c_I \det(\boldsymbol{\alpha}_I)\mathrm{Tr}\left[\boldsymbol{\alpha}_I^{-1}\frac{\partial \boldsymbol{\alpha}_I}{\partial T_{ij}}\right]. \tag{2.42}$$

It is important to note that the matrix $\partial \boldsymbol{\alpha}_I/\partial T_{ij}$ is either filled with zeros or with all zeros except for a single element with value one, depending on whether the element T_{ij} was one of those copied in to $\boldsymbol{\alpha}_I$ during its construction. Combined with the fact that $\boldsymbol{\alpha}_I^{-1}$ has already been evaluated for the formation of K_L, this sparsity means that only $\mathcal{O}(k_I^2)$ operations are needed for each configuration's contribution to the construction of $\partial \ln Q/\partial \boldsymbol{T}$ in this first stage of back propagation. To compare this with Filippi's approach, we note that Eq. (2.42) is equivalent to and under sparse evaluation has the same evaluation cost as Eq. (21) of Ref. [12].

For the kinetic energy, the chain rule in the first step of reverse accumulation gives

$$\frac{\partial K_L}{\partial \mu} = \frac{\partial t_0}{\partial \mu} + \mathrm{Tr}\left[\left(\frac{\partial K_L}{\partial \boldsymbol{T}}\right)^T \frac{\partial \boldsymbol{T}}{\partial \mu}\right] + \mathrm{Tr}\left[\left(\frac{\partial K_L}{\partial \boldsymbol{M}}\right)^T \frac{\partial \boldsymbol{M}}{\partial \mu}\right]. \tag{2.43}$$

Taking the last term first and noting that $\partial \boldsymbol{\beta}_I / \partial M_{ij}$ and $\partial \boldsymbol{\alpha}_I / \partial T_{ij}$ are equal as $\boldsymbol{\beta}_I$ is formed from $\boldsymbol{M}$ via the same copying pattern by which $\boldsymbol{\alpha}_I$ is formed from $\boldsymbol{T}$, we have

$$\frac{\partial K_L}{\partial M_{ij}} = \frac{1}{Q}\sum_{I=1}^{N} c_I \det(\boldsymbol{\alpha}_I)\mathrm{Tr}\left[\boldsymbol{\alpha}_I^{-1}\frac{\partial \boldsymbol{\beta}_I}{\partial M_{ij}}\right] = \frac{\partial \ln Q}{\partial T_{ij}} \tag{2.44}$$

which has already been evaluated when constructing the wave function derivatives and so incurs no additional cost. The middle term in Eq. (2.43) does require additional work, but by again using the determinant derivative and also the relationship

$$\frac{\partial}{\partial \boldsymbol{\alpha}_I}\mathrm{Tr}[\boldsymbol{\alpha}_I^{-1}\boldsymbol{\beta}_I] = -(\boldsymbol{\alpha}_I^{-1}\boldsymbol{\beta}_I\boldsymbol{\alpha}_I^{-1})^T \tag{2.45}$$

we can write

$$\frac{\partial K_L}{\partial T_{ij}} = (t_0 - K_L)\frac{\partial \ln Q}{\partial T_{ij}} + \sum_{I=1}^{N} c_I \det(\boldsymbol{\alpha}_I)\left(\mathrm{Tr}[\boldsymbol{\alpha}_I^{-1}\boldsymbol{\beta}_I]\mathrm{Tr}\left[\boldsymbol{\alpha}_I^{-1}\frac{\partial \boldsymbol{\alpha}_I}{\partial T_{ij}}\right] - \mathrm{Tr}\left[\boldsymbol{\alpha}_I^{-1}\boldsymbol{\beta}_I\boldsymbol{\alpha}_I^{-1}\frac{\partial \boldsymbol{\alpha}_I}{\partial T_{ij}}\right]\right) \tag{2.46}$$

whose evaluation incurs an additional cost that goes as $\mathcal{O}(k_I^3)$ per configuration after again exploiting the sparsity of the $\partial \boldsymbol{\alpha}_I / \partial T_{ij}$ derivatives. At this point we find a second important parallel between the AD approach and that of Filippi and coworkers, as Eq. (2.46) is equivalent to and has the same per-configuration cost as the configurational sum within Eq. (33) of Ref. [12]. With both $\partial K_L / \partial \boldsymbol{T}$ and $\partial K_L / \partial \boldsymbol{M}$ evaluated, reverse accumulation can now proceed through the back propagation graphs of $\boldsymbol{T}$ and $\boldsymbol{M}$ at a cost that is independent of the number of configurations. Again, it is crucial to point out here that regardless of how many and what type of parameters μ we are taking derivatives for, the derivatives $\partial \boldsymbol{T} / \partial \mu$ and $\partial \boldsymbol{M} / \partial \mu$ need not be explicitly constructed when performing reverse accumulation. In the case of orbital rotations or nuclear positions, the computational graphs for $\boldsymbol{T}$ and $\boldsymbol{M}$ are simple to back propagate through, and this last step will have essentially the same cost as their initial construction.

In summary, the direct application of AD to the wave function and local kinetic energy derivatives results in expensive terms (those that scale with the number of configurations N) that are entirely equivalent to the analogous terms in Filippi's approach. The details of the terms that do not scale with configuration number appear to be different, but in calculations with large numbers of configurations the cost of such terms is negligible. Thus, we find that AD offers no clear advantage over the methodology of Ref. [12], but it is nice to see it confirm the form of that methodology's most expensive and complicated components.

Simple Orbital Optimization Test

As an initial test of our orbital optimization implementation, we verify that it can remove the wave function's dependence on the initial orbitals in a low-symmetry, strongly correlated arrangement of four hydrogen atoms in which we construct a MSJ wave function for the ground state using the 10 most important configurations from a ground state CASSCF (4e,10o) calculation. With these configurations, we construct three different MSJ wave functions by employing molecular orbitals from RHF, B3LYP, and the (4e,10o) CASSCF. As seen in Figure 2.2A, ground state energy optimizations in which the orbitals are held fixed and only the Jastrow and CI coefficients varied result in three distinct energies, but when we then optimize the orbitals as well all three wave functions converge to the same energy, showing that the orbital optimization successfully removes the dependence on starting orbitals in this case.

Although the effects are not large, the orbital optimization does have a statistically significant effect on DMC energies, as Figure 2.2B shows the orbital-optimized nodal surface to be superior to that of any of the three wave functions in which only the Jastrow and CI coefficients were optimized.

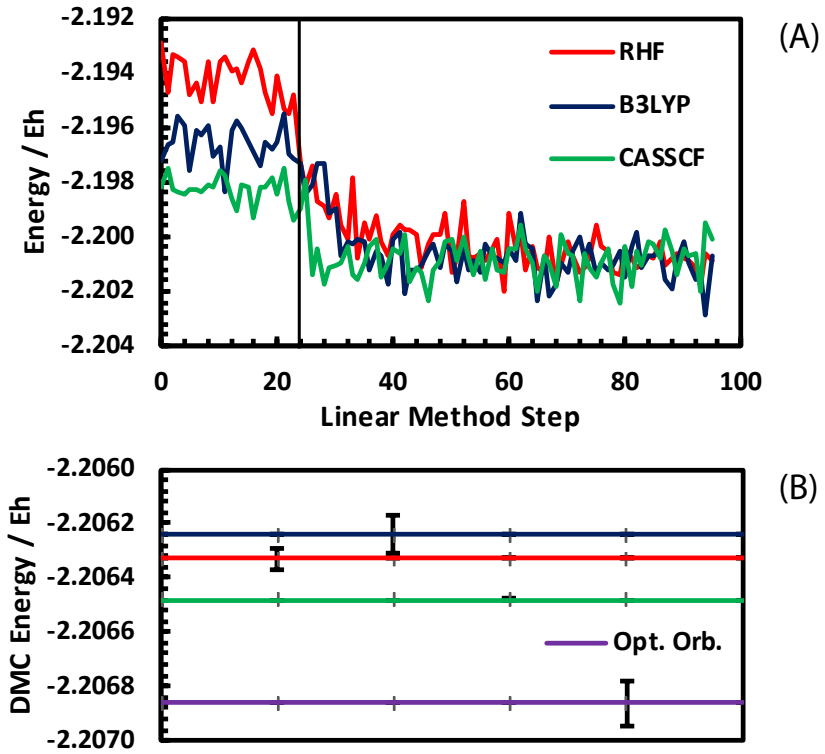


Figure 2.2: A: Energy vs optimization step for H_4 starting from different initial guesses for the orbitals. The orbitals are held fixed and only the CI coefficients and Jastrow variables optimized during the first 25 steps, after which all variables are optimized together. B: Time-step extrapolated DMC energies using the optimized trials wave functions before and after orbital optimization.

Chapter 3

Aperiodic systems

3.1 Formaldimine (CH_2NH)

Photoisomerization is an important phenomenon that is responsible for many interesting chemical events both in nature and the laboratory setting and is an excellent example of the intersection between multi-reference wave functions and excited states. Molecules that undergo photoisomerization include rhodopsins, retinal proteins involved in the conversion of light to electrical signals [122], and azobenzene, the prototypical photo switch studied for potential applications as a molecular motor. [136] One of the smallest molecules that undergoes photoisomerization is formaldimine (CH_2NH), in which the process proceeds following an absorption that promotes it to its lowest singlet excited state. [197, 52, 177] The subsequent rotation around the C=N bond mixes the σ and π orbitals and has been well studied, including by molecular dynamics simulations, [65, 202] and so this molecule makes for an excellent system in which to test the effects of state-specific orbital optimization and variance matching with modest MSJ expansions.

We model the ground and HOMO-LUMO ($n \rightarrow \pi^*$) excitation for torsion angles of 0, 45, and 90 degrees (an $A' \rightarrow A''$ transition for the 0° and 90° geometries where the molecule has C_s symmetry) using BFD effective core potentials and their VTZ basis. At each geometry, we ran a CIPSI calculation on the two lowest singlet states until it had accumulated at least 5,000 configurations. MSJ wave functions were then constructed by taking configurations with CI coefficients above a threshold (we tested three different thresholds: 0.02, 0.01, and 0.005) for both the ground and excited states. The resulting QMC calculations were then used to create our NLFFs to perform variance matching (an example of NLFFs is presented in Figure 3.1 for the 0° torsion coordinate). As twisting the C=N double bond introduces correlation, the number of configurations associated with a certain threshold varies at different torsion coordinates. For example, the ground state at a CI threshold of 0.01 corresponds to 5 configurations for the 0° geometry, 74 configurations for the 45° geometry, and 122 configurations for the 90° geometry.

The reader will likely have noticed that, for the 0° geometry presented in Figure 3.1,

we have reversed the role of ground and excited state in the variance matching procedure. This was done because of the unusual fact that, at this geometry, the ground state variances for the different CI thresholds were higher than the corresponding thresholds' excited state variances. Indeed, Figure 3.1 reveals that the lowest-threshold ground state variance that we used as the variance to be matched was higher even than the excited state's middle-threshold variance. In all cases we have studied in both this paper and elsewhere, this is the only one where we have observed the excited state having lower variances at a given threshold. Rather than changing the thresholds and hoping the situation would reverse itself, we instead simply reversed the role of ground and excited state in the variance matching in order to keep with the procedure's intention of fitting points on the lower-variance state to match the best wave function available for the higher-variance state.

For determining the rotational barrier heights, we took energy differences between the 90° and 0° geometries for both the ground S0 and excited S1 state. In this case, variance matching was performed separately for the S0 and S1 state barriers. As when computing excitation energies, we used the three different expansions coming from our three different thresholds to construct NLFFs for variance matching. As mentioned above, the ground state at the 0° geometry had a high variance compared to all other geometries and states. It was therefore most appropriate to use the lowest-threshold variance at the 0° geometry as the variance to be matched and to interpolate via NLFFs between variances of the different expansions at the 90° geometry, which we did separately for evaluating both the S0 and S1 barrier heights.

One interesting observation here is that the CI expansions derived from truncated CIPSI had as many as 57% of their configurations lying outside of a full valence (12e,12o) active space, which echoes previous studies in which selective CI wave functions often find many out-of-active-space configurations that prove to be more important than most of the active space configurations. That more than half of the 600 most important determinants in a CIPSI wave function lie outside of an active space that contains 853,776 determinants reminds us how significant this effect can be. It is also important to note that, thanks to starting from a truncated CIPSI wave function, our approach does not require selecting an active space in this molecule, which removes one of the more vexing difficulties of many multi-reference methods. In general this approach might be expected to lead to unaffordable selective CI calculations, but since we are truncating these wave functions to very modest CI expansion lengths anyways, we do not necessarily need CIPSI to converge to form a useful MSJ wave function out of it, a point to which we will return in Section 3.3.

As seen in Figure 3.2, our approach — which incorporates modest CI expansions, state-specific orbital optimization, and variance matching — predicts energy differences for excitations and barrier heights that are within 0.1 eV of full-valence active space (12e,12o) state-averaged MRCI+Q in all cases. It is especially noteworthy that the alternative approach of taking energy differences between fully optimized MSJ wave functions in which configurations for both states are selected via a shared CI coefficient threshold is much less reliable than the variance matching approach. This juxtaposition is a reminder that balancing wave function quality is crucial when working with unconverged CI expansions. Although

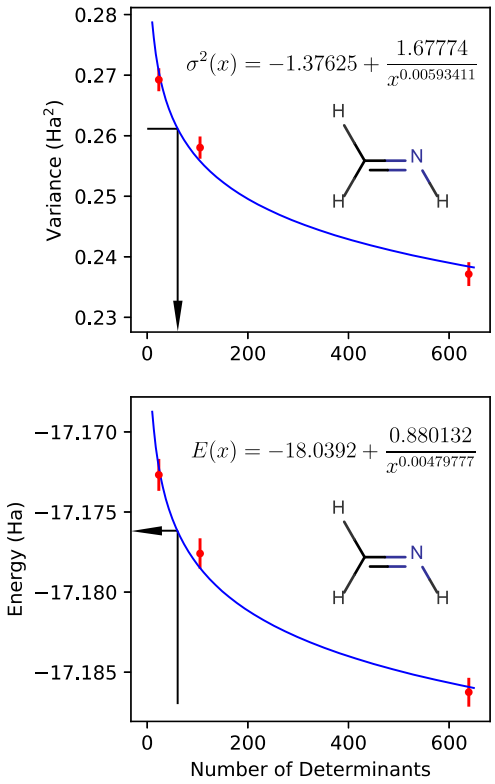


Figure 3.1: Plots of the NLFF fits from Eq. (2.14) (blue curves) for CH_2NH at the 0^o geometry. The excited state variance and energy data points being fit to are shown in red. The arrows demonstrate the variance matching procedure. In the top plot, the arrow takes the ground state variance and uses the NLFF to find the corresponding number of determinants the excited state would need to match that variance. In the bottom plot, the NLFF is used to convert this number into a ground state energy. Note that in this case (for reasons discussed in Section 3.1) we fit the excited state data and interpolate based on an input ground state variance, but the procedure is more typically applied by fitting the ground state data and interpolating based on an excited state variance.

this system is of course small enough that large brute force expansions are feasible (indeed CASPT2 also gives highly accurate results when used with a full-valence CAS), we emphasize that in large systems such an exhaustive approach will not be feasible and CI-based methods will be forced to work with incomplete CI expansions if they are to be used at all. The fact that our overall approach is able to be successful in this case despite using incomplete CI expansions is thus quite encouraging.

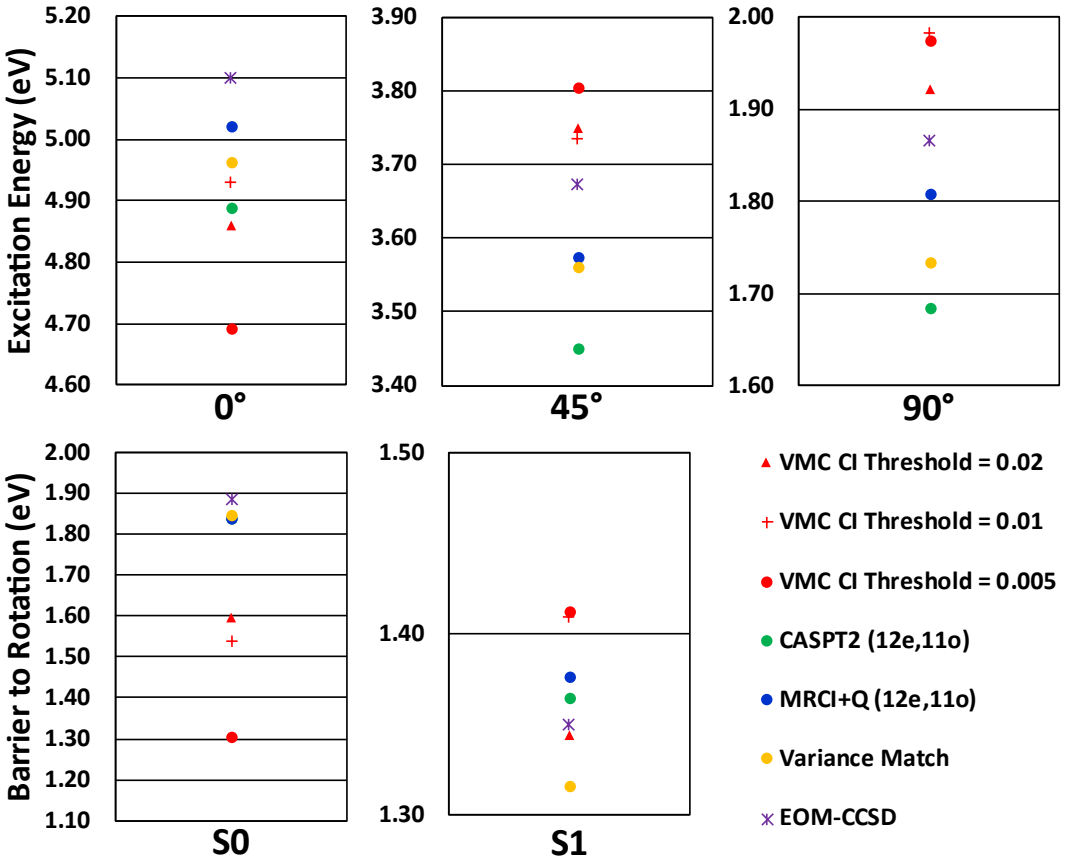


Figure 3.2: Top row: excitation energies for CH_2NH at various torsion angles. Bottom row: barriers to rotation on the ground state (S0) and excited state (S1) surfaces. Statistical uncertainties were less than 0.04 ev in all cases

3.2 Thioformaldehyde (SCH_2)

Although CH_2S undergoes similar chemical reactions as CH_2O [171, 43] and sees a similar change (about 0.8 Debeye) in its MRCI dipole moment during its low-lying $n \rightarrow \pi^*$ singlet excitation (an $A_1 \rightarrow A_2$ transition), this absorption band is red-shifted so that what lay in the near ultraviolet in CH_2O lies in the visible region [43] for CH_2S. Given sulphur's more labile valence electrons and the persistence of modest charge transfer character, CH_2S makes for an interesting test case, especially because exact results can be benchmarked against even in a triple zeta basis by employing large-core pseudopotentials so that only 12 electrons need to be simulated explicitly.

We took the approach described in the theory sections in order to try to ensure balanced MSJ descriptions of the two states. Specifically, we optimized the orbitals, CI coefficients,

and Jastrow variables for an 875 determinant MSJ expansion for the first excited singlet state using determinants drawn from a two-state CIPSI calculation in the RHF orbital basis. We then performed a series of analogous ground state optimizations using expansions with 2, 5, 10, 50, 100, and 875 determinants taken from the same CIPSI calculation to perform our NLFF. After fitting our NLFF to the resulting energies and variances, our variance-matched approach predicts an excitation energy similar to that of full-valence state-specific CASPT2, as seen in Table 3.1. The correct excitation energy in this basis — confirmed by the agreement of MRCI+Q with extrapolated SHCI — is about 0.2 eV higher. We therefore see that, although our approach to balancing the accuracies of the different states' descriptions is not perfect, it is able to provide reasonably high accuracy with very short CI expansions.

A feature of CH_2S that is worth noting is that our multi-reference quantum chemistry results are quite insensitive to whether we a) ignore the molecule's symmetry and arrive at the ground and first excited singlets via a 2-state state average or b) exploit the molecule's symmetry in order to treat both the ground and excited state as ground states of their respective symmetry representations. Table 3.1 shows that the CASSCF, CASPT2, and MRCI+Q excitation energies are little changed when we switch between these state-averaged and state-specific approaches. Certainly our ability to afford a full-valence CAS in thioformaldehyde contributes to this insensitivity, but in any case our MSJ approach's ability to tailor the orbitals in a state-specific manner is clearly not essential here. We now turn to a case in which state averaging is more problematic in order to emphasize the advantages of a fully variational approach with state-specific orbitals.

Table 3.1: Excitation energies for the lowest singlet excitation in thioformaldehyde. CASSCF, CASPT2, and MRCI+Q used a full-valence (12e,10o) active space, while SHCI (and the CIPSI calculation from which we generated the MSJ expansion) was performed for all 12 electrons in all orbitals.

Method	hν / eV
2-state-SA-CASSCF	2.68
SS-CASSCF	2.65
2-state-SA-CASPT2	2.16
SS-CASPT2	2.13
2-state-SA-MRCI+Q	2.31
SS-MRCI+Q	2.32
SHCI	2.31(1)
EOM-CCSD	2.40
Variance Matched VMC	2.07(2)

3.3 $\left[C_3N_2O_2H_4Cl\right]^-$

Compared to CH_2S, many of the low-lying excited states of $\left[C_3N_2O_2H_4Cl\right]^-$, shown in Figure 3.3, have very strong charge transfer character. In the ground state, this system localizes the extra electron on the Cl atom, which although not bonded covalently to the main molecule is attracted by dipole/charge interactions and dispersion forces (we determined its location through MP2/cc-pVDZ geometry optimization). According to EOM-CCSD, the first four singlet excited states all transfer an electron into the lowest π^* orbital. In order of increasing excitation energy, these transfers come from the two Cl in-plane p orbitals (3.56 and 3.74 eV), the Cl out-of-plane p orbital (3.86 eV), and an oxygen in-plane p orbital (3.91 eV). In addition, there are multiple other $n \rightarrow \pi^*$ and $\pi \rightarrow \pi^*$ singlet transitions in the 4−7 eV range. Although some of these states have strong charge transfer character, EOM-CCSD is a good reference for their excitation energies thanks to: (a) the fact that they are all singly excited states and (b) EOM-CCSD's doubles operator's ability to provide the state-specific orbital relaxations that are so crucial for charge transfer.

In contrast, multi-reference methods are harder to use effectively here, both due to the molecule's larger size and due to the difficulties that state averaging encounters when faced with states that have large differences between their charge distributions. Most notably, the ground state and any excited states that do not involve the chlorine atom have dipoles that differ by more than 10 Debeye compared to excited states that have transferred an electron from the chlorine into the π network. These differences mean that even orbitals outside the active space are expected to relax significantly when transferring between these two sets of states, making it exceedingly challenging to arrive at a good set of state averaged orbitals that are appropriate for all of the low-lying states. This difficulty can be seen even if we

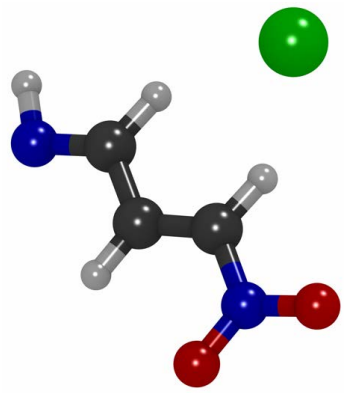


Figure 3.3: The $\left[C_3N_2O_2H_4Cl\right]^-$ anion used in our chlorine-to-π^* charge transfer example. In the ground state, the charge is localized on the (green) chlorine atom, while many of the excited states have the charge distributed in the π network.

Table 3.2: Excitation energies in eV for the totally symmetric chlorine-to-π^* transition in $\left[C_3N_2O_2H_4Cl\right]^-$ using BFD effective core potentials and the corresponding VDZ basis set. [29] CASSCF and CASPT2 results are based on either state averaged (SA) orbitals or nearly state specific (SS) orbitals. CASPT2 employed either Roos-Andersson (RA) level shifts [164] or the ionization potential electron affinity (IPEA) approach [69] to deal with intruder states. See Section 3.3 for details. VM-VMC (g,e) stands for Variance Matched VMC with g/e being an integer value representing the number of determinants used in the ground (g) and excited (e) state wave functions.

Method	Level Shift		hν
4-state-SA-CASSCF	N/A		2.127
2-state-SA-CASSCF	N/A		3.980
(95/5)-SS-CASSCF	N/A		4.669
4-state-SA-CASPT2	RA	$\varepsilon = 0.2$	3.369
4-state-SA-CASPT2	RA	$\varepsilon = 0.3$	3.340
4-state-SA-CASPT2	IPEA	$\varepsilon = 0.25$	3.461
2-state-SA-CASPT2	RA	$\varepsilon = 0.2$	3.285
2-state-SA-CASPT2	RA	$\varepsilon = 0.3$	3.304
2-state-SA-CASPT2	IPEA	$\varepsilon = 0.25$	3.546
(95/5)-SS-CASPT2	RA	$\varepsilon = 0.2$	3.365
(95/5)-SS-CASPT2	RA	$\varepsilon = 0.3$	3.387
(95/5)-SS-CASPT2	IPEA	$\varepsilon = 0.25$	3.541
PBE0	N/A		1.233
B3LYP	N/A		0.903
M06-2X	N/A		0.903
wB97X-V	N/A		3.348
wB97M-V	N/A		3.187
EOM-CCSD	N/A		3.856
VM-VMC(10,200)	N/A		3.80(3)
VM-VMC(99,1000)	N/A		3.87(1)

restrict our attention to the two lowest states in the totally symmetric representation, which are the ground state (charge on the chlorine) and the out-of-plane-Cl-3p $\rightarrow \pi^*$ excitation (a totally symmetric $A' \rightarrow A'$ transition). As seen in Table 3.2, there is a 0.7 eV difference in the excitation energy predicted by an equally-weighted 2-state state averaged CASSCF calculation and a calculation in which we attempt to optimize the orbitals state specifically by using 95%/5% and 5%/95% weightings (note that our active space distributed ten electrons among the three chlorine 3p orbitals and the five π/π^* orbitals closest to the gap). The CASSCF sensitivity to state averaging is thus much higher here than it was in CH_2S, and, as we discuss in the Supporting Information and show in the 4-state-SA-CASSCF results of Table 3.2, the problem would be worse-still if the molecule's symmetry plane were not present.

As real chemical environments such as protein superstructures or solvents typically remove symmetry, one realizes that these types of state averaging difficulties will be quite common when attempting to model charge transfer in large molecules and realistic environments.

Of course, one should not expect quantitative accuracy from CASSCF excitation energies whether or not there are state averaging concerns, as these calculations omit the weak correlation effects of orbitals and electrons outside the active space. For larger molecules and active spaces, CASPT2 is much more affordable than MRCI+Q, and so we have employed it here both to include weak correlation effects and in the hope that it can via its singles excitations help to put back the state-specific orbital relaxations that are inevitably compromised during state averaging. Unfortunately, we found that the excited state we are after suffers from intruder state [164] problems in all cases here, regardless of how the state-averaging was handled, and so we were forced to employ level shifts in order to avoid unphysically large perturbative corrections. We found that for each of our state-averaging and state-specific approaches, the value and type of level shift made a noticeable difference in the CASPT2 excitation energies, which is not ideal. Overall, the CASPT2 results's errors ranged from about 0.3 to 0.5 eV when compared to EOM-CCSD, which does leave something to be desired but is nonetheless an improvement over TD-DFT.

For our MSJ treatment of this system, we began by iterating the variational stage of a 4 state CISPI calculation (ignoring symmetry) in the RHF orbital basis until it had identified more than 5,000 important determinants. At this point we found that the out-of-plane charge transfer state we are focusing on was the fourth CIPSI root and we ended the CIPSI iterations, even though for a system of this size the expansion procedure was certainly far from converged. This incomplete CIPSI does not present an issue for us, though, as we imported only the fourth root's 200 most important determinants into our MSJ wave function, and we expect that by the time CIPSI has reached 5,000 determinants the identity of its leading 200 will be well established. We then applied our variational excited state methodology to optimize the orbitals, CI coefficients, and Jastrow variables for this MSJ wave function and evaluated its variance. Repeating this procedure for the ground state (the first root from the same incomplete CIPSI expansion) we found that its MSJ wave function required just 10 determinants in order to match the excited state variance. Given that the ground state variance will be a much more noticeably discreet function of determinant number for such short expansions, we decided to forgo the NLFF (which makes more sense when the variance is changing close to continuously with determinant number) and instead varied the number of ground state determinants by hand to find the expansion whose variance most closely matched that of the excited state. This approach found that the 10-determinant ground state MSJ wave function (with optimized orbitals, CI, and Jastrow) made for the best match, which resulted in a predicted excitation energy within 0.1 eV of the EOM-CCSD benchmark, as seen in Table 3.2. To test the robustness of this estimate, we repeated the procedure using the fourth root's 1,000 most important CIPSI determinants to construct a larger excited state wavefunction and in turn found that a 99 determinant ground state matched its variance. Using the larger wave functions changed the variance matching estimate for the excitation energy by less than 0.1 eV and moved it even closer

to the EOM-CCSD benchmark number. Thus, as in the smaller systems, we find that a combination of short CIPSI-derived expansions, orbital optimization, and variance matching delivers a reasonably high accuracy even in a case whose size and strong charge transfer character complicates the application of traditional multi-reference methods.

Chapter 4

Periodic Systems

4.1 Finite size error and explicit variance matching

In this section we describe the finite size error correction used for periodic system, not needed in our aperiodic systems. Then we describe in detail the procedure to calculate optical band gaps using explicit variance matching.

To address the finite size errors of our simulation cell for MgO and PAE we performed VMC-CIS followed by orbital optimization (which we define as VMC-CIS-OO) calculations with supercells containing N, 2N, and 4N atoms (in which N is the number of atoms in the minimal supercell that contains the optical transition of interest) and then performed an extrapolation for the predicted gap.[228] For MnO this type of finite size effect is too costly to perform because the minimal supercell that contains the optical transition is very large. Instead, for MnO, we perform a DMC finite size effect correction based on simple single-Slater wave functions for the ground and excited state similar to that done by others. [228, 111, 175]

The steps taken in the prediction of the optical gap are as follows ...

1. Construct and optimize a SJ wave function with chosen SPO set for the ground state.
2. Perform a VMC-CIS or VMC-CIS(D) calculation (see section 2.3) to determine initial excited state configuration weights
3. Optimize Jastrow, and Orbital coefficients for the ground state. Then Optimize Jastrow, CI, and Orbital coefficients for the excited state.
4. Perform explicit variance matching (see section 2.4) by varying the number of configurations in the excited state. If necessary repeat the last two steps for the excited state with an active space that is larger or smaller.
5. Repeat all previous steps for different supercell sizes. Then use data points to perform finite size correction of the optical gap via linear extrapolation.

The details of chosen pseudopotentials, SPO set, supercells sizes, matched variances, number of determinants in MSJ, finite size extrapolation, and DMC time step extrapolation can all be found in the Supplemental Material.

4.2 Magnesium Oxide (MgO)

Magnesium Oxide is a simple rocksalt structure material with a lattice constant of 4.21 Å.[225] The conventional picture of the material's band diagram is that the valence band is dominated by O 2p character and the conduction band is of Mg 3s character. The optical excitation can then be interpreted as causing an electron donation from O 2p $\rightarrow$ Mn 3s. It has been shown that standard single particle theories are quite adequate at predicting the optical gap of MgO.[26, 181] Although these single particles theories do not describe the essential electron-hole interaction of an optical absorption, the singlet excited state of MgO has only a small exciton binding energy [181, 16] making the approximation of noninteracting electron and hole acceptable.

Method	gap (eV)
Expt [1]	7.83(2)
LDA [2]	4.98
HF [3]	8.9
B3LYP [4]	7.6
G_0W_0 [5]	7.7
VMC-CIS-OO [6]	7.7(2)

Table 4.1: The optical band gap of MgO determined by different methods.

Table 4.1 summarizes prediction of various methods. As is commonly known LDA is underestimating the gap while HF overestimates the gap. The G_0W_0 approach under the random phase approximation taken by Schönberger and Aryasetiawan correctly increases the LDA gap by including dynamically screened Coulomb potential (which is known to work well for s-p systems such as MgO).[181]

Our approach to the optical band gap is also an improvement compared to the single particle theories and produces a result on par with the G_0W_0 approximation. We constructed our MSJ wave function with HSE06 orbitals and given the success of B3LYP at correctly describing MgO[26] are not surprised at the success of our approach. With qualitatively correct quasiparticle states our QMC orbital optimization does not need to overcome large barriers to produce a reasonable estimate to the optical band gap. See section 4.1 and B.1 for details of our VMC-CIS-OO calculation.

4.3 Trans-Polyacetylene (PAE)

In 2000 the Nobel prize in chemistry was awarded to Alan J. Heeger, Alan MacDiarmid and Hideki Shirakawa[203] for their 1977 work synthesizing conductive iodine-doped Polyacetylene (PAE).[193] Conducting and insulating polymers continue to be of interest due to their promising applications as organic light emitting diodes (O-LED) and subsequent use in display technologies.[31, 66]

Despite PAE being a very simple polymer, its insulating nature is not so simple. If one were to rely on elementary models for a one dimensional system such a "particle in a box" and "Hükel method" the HOMO-LUMO gap would vanish as the system size continued to increase which suggests metal behavior. But experimentally there is an observed insulating behavior with an optical gap of 1.5 eV.[113] By considering Peierl's theorem one can rationalize that the polymer chain undergoes distortion to produce a alternating short-long-short bond length and produces a band gap (analogous to a Jahn-Teller distortion lifting the degeneracy of electronic states).[149] In fact, this Peierl's distortion has been demonstrated experimentally in pristine PAE films[35]

While the Peierls distortion allows for a qualitatively correct prediction of a band gap there are other effects that still must be considered in order to produce a quantitative prediction of the optical gap. One must consider that the system is not truly one-dimensional but instead exists in a three-dimensional film. There are also of course the correlation effects between electrons, the interaction of π molecular orbitals on different chains, and the electron-hole binding effects to consider as well.

Because DFT/G_0W_0 calculations are intended for fundamental band gap predictions, these calculation can only serve as a first order approximation for the optical gap of PAE. DFT is known to underestimate band gaps, and while G_0W_0 calculation can be used to predict the fundamental gap well, it does not take into account an exciton binding energy necessary for an optical gap prediction. Theoretical prediction for the binding energy ranges from approximately 0.1 eV to 1.0 eV.[163, 39, 105, 90] The possibly large exciton binding energy necessitates solving the Beth-Salpeter equations to produce accurate results.[163, 204]. Furthermore comparing single chain to crystal studies of PAE reveal that interchain interactions are responsible for a reduction of both the electronic energy gap (relative to the G_0W_0 fundamental gap) and the binding energy of bound excitons.[204] [163]

In this work we study the lowest singlet transition of PAE using VMC-CIS-OO. We use the experimentally determined structure by Shimamura and coworkers.[192] Their electron diffraction work led to the assignment of an orthorhombic unit cell structure which contains two monomer units and belongs to the P_{nam} space group.[192]

For the construction of our wave function we use HF orbitals instead of HSE06 as a starting point because a VMC-CIS calculation resulted in CI vectors that were easy to interpret for HF orbitals while that of HSE06 were not. But, as can be seen in Table 4.2, the HF orbitals on their own vastly overestimate the gap. By performing parameter optimization of the Jastrow and orbital parameters in QMC we allow the orbitals to relax and accommodate the exciton and ultimately produce a reasonable optical gap of 1.38(2)

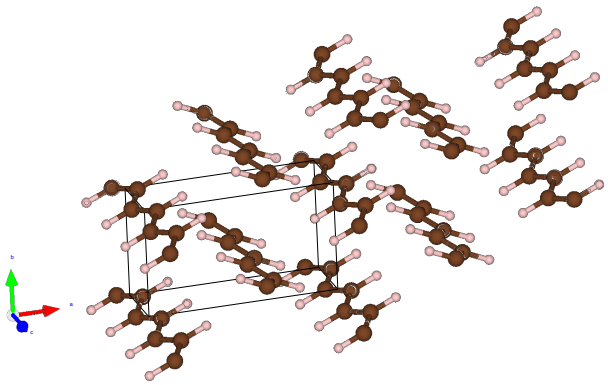


Figure 4.1: Packing arrangement of PAE represented in a 2x2x3 supercell.

eV. A value that is close to the experimental 1.5 eV gap [113] and on par with Beth-Salpeter results of 1.65 eV[204]. See section 4.1 and B.1 for details of our VMC-CIS-OO calculation.

Method	gap (eV)
Expt [113]	1.5
PBE	0.02
HF	16.70
HSE06	0.53
G_0W_0 [204]	1.8
G_0W_0/BSE [204]	1.65
VMC-CIS-OO	1.38(2)

Table 4.2: The optical band gap of PAE determined by different methods.

4.4 Iron Oxide (FeO)

Iron Oxide (a cubic, anti-ferromagnetic, transition metal oxide, bulk system) is a challenge for present day electronic structure theory. Its diverse structural, electronic, and magnetic transformations across a range of pressures and temperatures make it a difficult system to study both experimentally and theoretically. [32] But even at ambient temperature and pressure functions, such as the ability to reduce water and create hydrogen gas upon irradiation, are difficult to theoretically model. [123] A proper calculation of Iron Oxide's electronic structure need to include: the electron-electron interactions in d orbital subshells, the d-p orbital hybridization, and the charge transfer character exhibited in excited states, which

is just plain difficult to do simultaneously. [123] In fact, current theoretical methods fail to accurately predict its electronic, magnetic and structural properties simultaneously, and rely on empirical parameters. [123]

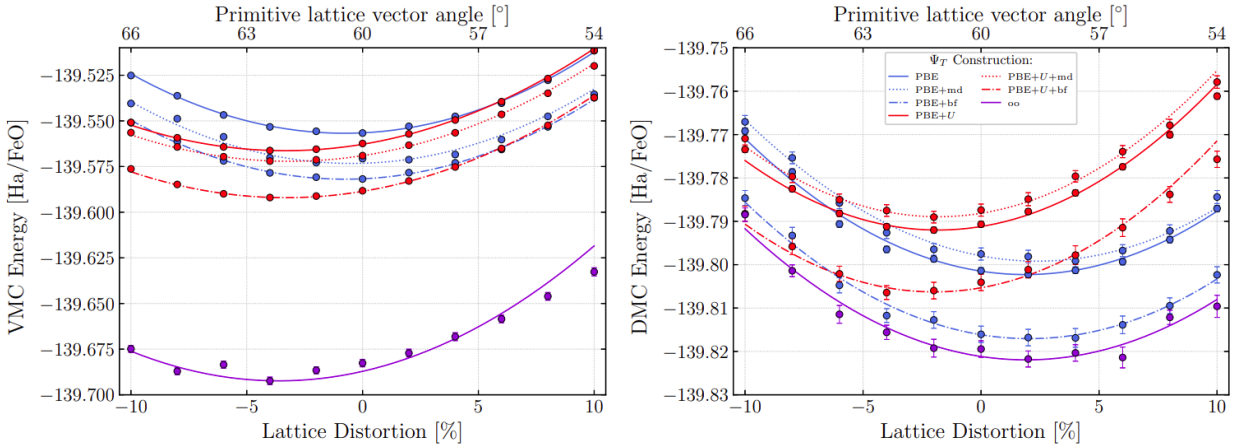


Figure 4.2: VMC (left) / DMC (right) strain curves constructed from various QMC wave functions.[210] The solid lines represent quadratic polynomial fits to the data points to guide the eye. The grey arrow indicates the experimentally determined lattice distortion.

For the study of FeO we demonstrate that Slater-Jastrow wave functions with orbital optimization can produce orbital starting point independent DMC energies, spin-density, and equilibrium lattice distortion that coincide with experimental values. Furthermore the simple Slater-Jastrow wave function with orbital optimization outperforms traditionally more sophisticated wave functions regardless of whether the SPOSet was PBE or PBE+U (U = 4.3 eV).

Figure 4.2 illustrates the various strain curves determined by different wave functions. Details of the all wave functions calculations can be found in (Townsend 2020 [210]). In the VMC plot it appears that the minimum of the strain curve produced is highly dependent on the chosen SPO set but less so on the type of QMC wave function used. PBE+U orbitals produce a negative strain minimum while PBE orbitals produce a positive strain minimum. Considering a set of wave functions using the same SPO set the ordering of the strain curves with respect to energy from highest to lowest is SJ, MSJ, BF whether PBE or PBE+U orbitlas are used. Now considering the DMC plot, we see that the minimum strain of the PBE orbitals is now positive, and the minimum strain of the PBE+U curves is now negative (the opposite of the VMC plot). Again, considering the wave function with a common SPO set the order from highest energy to lowest is now MSJ, SJ and BF.

Now we consider the wave function that is PBE+SJ+OO referred to simply as 'OO'. The OO curve in the VMC plot has a significant improvement in the energy over any other wave function by about 0.1 Eh. Although, despite the variational improvement of the energy the

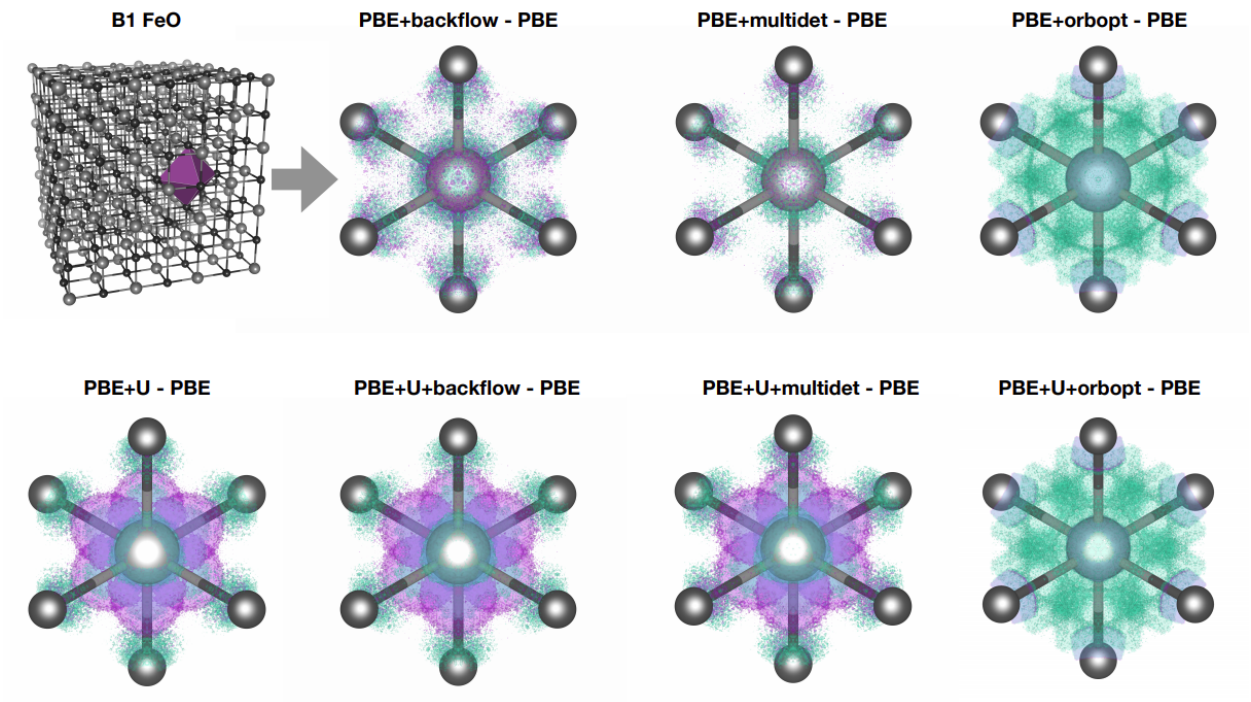


Figure 4.3: Spin density plots of FeO using various types of QMC wave functions.[210]

minimum of the strain curve is still predicted to be negative. But the OO curve in the DMC plot correctly predicts a positive strain minimum and again is the lowest curve with respect to energy.

With the data presented so far one would like to conclude that orbital optimization is drastically correcting the nodal surface of the wave function to produce correct DMC result regardless of starting orbital set. Although the OO performance in the VMC plot casts doubt to whether this is actually true. Comparing the spin density plots of the various SJ wave functions with different starting orbitals would help clarify if the nodal surface is truly being transformed consistently. Figure 4.3 plots the difference in density of the specified wave function with respect to a PBE+SJ wave function. The plots illustrate how the density for PBE+U wave functions increases at the Fe centers (as is expected for a +U correction of transition metal oxide systems). The PBE+MSJ and PBE+BF wave functions seem to change the spin density very little. Now the most significant finding is that PBE+SJ+OO and PBE+U+SJ+OO both produce vary similar spin density plots. This suggests that indeed that the orbitals are being robustly optimized.

The most important consequence of this study is not that the minimum of the strain curve is correctly produced but that it was arrived at in a way that was parameter free (by avoiding the need of using +U correction or a particular functional). We discuss the

motivations and consequences of using the +U correction later in section 4.5 when exploring the optical gap of the bulk MnO system. But for now suffice it to say that the even if PBE+U produces qualitatively correct SPO description of the state there is a question of how to select a value of the U parameter that will lead to reliable quantitative description of the system. Therefore this 'parameter free' QMC approach can be invaluable for systems in which chemical intution may not be available to include corrections like +U.

4.5 Manganese Oxide (MnO)

Although decades of work have been performed on Transition metal oxides (TMOs) a truly *ab initio* prediction of the gap and the nature of the excitation remains elusive. TMOs have been of theoretical interest as far back as 1937, when single particle theory (band theory) predicted metallic behavior while experimental evidence revealed an insulating nature.[21] In this section we take particular interest in the $\Gamma \rightarrow \Gamma$ and $Z \rightarrow \Gamma$ optical transitions of MnO, a TMO with partially filled d-orbitals.

MnO crystallizes into an almost rock-salt (cubic) structure with a small rhombohedral distortion of $\alpha = 90.624(8)^\circ$ and a lattice constant of 4.4316(3) Å.[129, 170] Below it's Neel temperature of 116K the system transforms into an AF2 phase.[129, 170] Although decades of work have been performed on TMOs a truly *ab initio* prediction of MnO's optical gap remains elusive. In addition whether its optical excitation is "charge transfer" or Mott-insulator type is still unclear theoretically and experimentally. [111, 223, 184, 224]

Past progress on optical gap prediction

Through the 1940s-1960s work on model systems, that included strong on site coulombic repulsion, by Mott and Hubbard laid the conceptual foundation for how we presently think of TMOs as Mott-insulators that can be understood via the Hubbard model.[132, 92] These models helped scientist understand that the failure of band theory and DFT is due to the inability of these theories to describe the on site effects of strongly localized and correlated d-electrons of the transition metal.

Brandow created a qualitative picture of the electronic structure of MnO by considering the crystal field theory splitting of the transition metal d-orbitals surrounded by oxygens in an octehedral geometry, and then defining a Hubbard like Hamiltonian with parameters for different d-orbitals interactions. These Hubbard parameters included U (local coulombic interaction of two opposite spin electrons occupying the same d-oribital), U' (local coulombic interaction of two opposite spin electrons occupying different d-oribitals), and J (local exchange interaction of two parallel spin electrons occupying different d-orbitals).[25] Using spectroscopic data to determine the values of these parameters Brandow could make semi-quantitative assements of many features of NiO and CoO.[25]

With the insight provided by Brandow it was shown in 1994 that Unrestricted Hartree Fock (UHF) can qualitatively produce the correct ground state for NiO and MnO.[209]

Towler and others explain that the fact that UHF correctly captures exact exchange eliminates the self interaction error present in DFT but also reproduces an 'orbital occupation dependent potential' (which is generally referred to in the literature as 'orbital polarization') that Brandow produced via his Hubbard-Hamiltonian.[209]

Although UHF produces an adequate ground state picture of TMOs it lacks correlation effects and therefore it was desirable to find a way to adjust the approximations of DFT in order to capture the 'orbital polarization'. In 1997 the LDA+U method was introduced, which takes into account orbital dependence of the Coulomb and exchange interactions (absent in LDA) through a Hubbard like U parameter, and gives a much improved description of some Mott insulators.[11, 44]

Throughout the 1990s to early 2000s hybrid functionals for DFT (such as B3LYP, PBE0 and HSE) that included a fraction of exact-exchange were introduced.[15, 150, 87] These functionals mitigated the self interaction error, which allow for d-orbitals to become more localize and help reproduce more correct band gaps.[68]

The progressive success of every theory/model that has been summarized has depended on a approximate description of orbital polarization to produce quality single particle orbital (SPO) sets. The one particle Green's function can be constructed from these SPO sets under the assumption that the DFT Kohn-Sham states are good approximations to the quasiparticle states. If the best SPO sets produced are similar enough to each other then subsequent G_0W_0 calculations should be able to add the many body effects necessary to correctly predict a consistent insulating gap.

Throughout the 2000s to present day these G_0W_0 calculation have been attempted but a wide range of band gaps that depend on the SPO set utilized resulted.[174, 100] Creating correct orbitals to describe TMOs has proven a difficult task and has been accomplished by introducing empirical approximations (+U and mixing exact exchange). It is unfortunate that the best SPO sets from these approximation do not produce consistent G_0W_0 results, and even self consistent versions of the calculation (GW_0) have strong starting point dependence on the SPO set.[174] For example, Jiang and coworkers found that varying the value of the 'U' parameter from 0-8 eV led to predicted G_0W_0 gaps in the range of $\approx 1.5 - 8$ eV for MnO.[100]

In 2004 Wagner and coworkers investigated the optical gap of MnO via DMC and made comparisons among various SPO sets.[111] They built the simplest possible trial wave functions (single Slater Jastrow) consisting of the Aufbau configuration for the ground state, and only a single particle-hole excitation for the excited state. These wave functions were constructed using orbitals from UHF, and the DFT functionals PW86 and B3LYP. They were able to evaluate the cohesive and excitation energies of MnO and found that using Ne-core pseudopotentials for Mn was important. In their comparison of various SPO sets from UHF, PW86, and B3LYP they also showed that the fraction of exact exchange present in the exchange correlation functional dictated the degree of s-p hybridization.

In Wagner and coworkers more recent 2015 work with MnO they used the "PBE1 x " framework, in which the degree of exact exchange mixing is determined by parameter α, to again show that α was indeed essentially a tuning parameter for hybridization. [180]

Because the FN-DMC description of the the optimal phase and lattice constant of MnO differed from the DFT description an analysis of site-resolved charge fluctuations, also known as the compressibility, was also done.[180] The result of the analysis showed that DMC tended to localize charge when compared to PBE, PBE0, HSE06 but delocalized the charge compared to HF. Using the optimal α parameter for the ground state they also estimated the fundamental gap and optical gap of MnO via FN-DMC using SJ wave functions but ultimately produced overestimates when compared to experiments.[180]

Past insight into the nature of optical gap

Typical analysis of the type of transition for MnO begins with the 'ionic model' of the system. That is, when considering the bulk MnO system ground state we can interpret the electrons of the Mn 4s-orbital to be donated to an O 2p-orbital which creates the ionic ground state picture of Mn^{+2} O^{-2}. [224, 184, 223, 25] Next the crystal field theory splitting of the d-orbtials centered at the Mn atom is considered. Due to the octahedral arrangement of oxygen atoms the center Mn atom's d-orbitals split into two higher e_g states ($d_{x^2-y^2}$ and d_{z^2}) three lower t_{2g} states (d_{xy}, d_{xz}, d_{yz}). It is expected that the occupied d-orbitals associated with the transition metal are localized and do not overlap much with neighboring atoms due to the high Z value of Mn.[25]

From the current picture described there are four main transitions of the band gap of MnO that have been discussed in the literature, the two most popular being "Mott Hubbard" and "charge transfer". [224, 184, 223, 25]

- TM d $\rightarrow$ TM d (Mott Hubbard)
- TM d $\rightarrow$ TM 4s
- anion p $\rightarrow$ TM d (charge transfer)
- anion p $\rightarrow$ TM 4s

Although there have been much attention to the "Mott Hubbard" and "charge transfer" transitions there has been less so for the remaining two.

Traditional analysis from this point continues by one of two ways. The first is building a model Hamiltonian with parameters that are either rationalized or extracted from experimental findings. Then there is verification of the model Hamiltonian via experimental trends. [224, 184, 223, 25] The second is performing an *ab initio* calculation and then reconciling its result with one of the mentioned descriptions of the gap. In this work we attempt the second route of performing an *ab initio* calculation and reconciling its results with one of the above descriptions of the gap. It should be noted that previous QMC studies of MnO simply performed DMC using SJ wfn while we will be attempting to use more sophisticated methods.

QMC wave function creation

Our VMC-CIS(D) with orbital optimization method (referred to as VMC-CIS(D)-OO) method allows us to work with the correlated many-body wave function for excited states directly and we use it to predict the optical gap of MnO. The Jastrow factor will take into account the dynamic correlation associated with interelectronic cusps, and nuclear cusps directly. Relaxation and screening effects that are missing in single particle theories are then corrected for by using orbital optimization to variationally improve the excited state. Finally if strong correlation is present our VMC-CIS(D) method for configuration selection could capture the necessary configurations. See section 4.1 and B.1 for details of our calculations.

Choice of active space

By considering all the configuration that correspond to the transition types mentioned in Section 4.5 in our VMC-CIS-OO wave function we can determine the nature of the excitation. We therefore performed a VMC-CIS(D) calculation in an active space that includes the orbitals relevant for the determination of "charge transfer" / "Mott Hubbard" / "other" nature. These orbitals include the set of O 2p, O 3s, Mn 3d, and Mn 4s of all atoms in the unit cell.

Choice of SPO set

Now that an active space is chosen, we simply need to decide on a starting orbital set for the wave function. Traditionally QMC methods must use and is limited by an input SPO set just as G_0W_0 calculations do and are.[63, 142, 185] It is therefore worth highlighting the shortcoming of the various SPO sets to explain why G_0W_0 calculations fall short, and to justify the decision for an input SPO set for our work.

Hartree-Fock orbitals are known to overestimate the band gap of materials due to the virtual orbitals experiencing the mean potential produced by the N occupied-orbitals as opposed to correctly experiencing the N-1 electrons that should be felt by a neutral excitation (and of course no correlation effects are included). These shortcomings cause UHF to produce a band gap of 12.9 eV for MnO, which is over three times the experimental value.[209]

DFT orbitals are known to underestimate the band gap due to the self-interaction error. LDA incorrectly cannot split the energies of the d-orbitals due to how the functional takes into account the effect of exchange in an average manner.[209] Although hybrid functionals and the +U approximation ameliorate the issue they present their own problems in ambiguity of how to be parameterized. [88] For hybrid functions this can be as simple as what percentage of exact exchange to use. For +U approximations there is the question of how to determine the U parameter unbiasedly and how to account for the double counting error of the energy.[88] But even without these errors, by assuming the exact exchange correlational functional was known, the Kohn-Sham band structure does not provide the fundamental band gap of the real interacting-electron system as it does not include the finite and posi-

tive derivative discontinuity of the Exchange-Correlation energy as a function of number of particles. [153, 151, 187, 172]

Now that we have compared the SPO set choices we must make a selection. But before making a selection we show here that with QMC orbital optimization the dependency of the starting orbitals can be lifted (assuming that optimization is robust and the basis is sufficiently 'complete'). For the minimal 4-atom magnetic unit cell we summarize the resultant ground state energy for a single Slater wave function that uses either LDA or LDA+U orbitals and one/two body Jastrow factors in Table 4.3. We see a difference of 59 mEh between the two wave functions for the VMC energy and a difference of 8 mEh for resultant DMC energies.

Method	LDA	LDA+U
VMC	-240.065(1)	-240.124(1)
DMC	-240.504(1)	-240.512(1)
VMC-OO	-240.175(1)	-240.175(1)
DMC	-240.518(1)	-240.517(1)

Table 4.3: Table of energies from various QMC wave functions that include the one and two body Jastrow factor and an orbital basis determined from LDA or LDA+U DFT calculations. The parameter values of U = 4.7 eV and J = 0.8 eV values were obtained from constrained DFT calculations determined by Jiang and co-workers [100]

Table 4.3 also shows resultant VMC / DMC energies after QMC orbital optimization has been performed. The VMC and DMC energies of the LDA wave function improved by 110 mEh and 14 mEh respectively. While the LDA+U VMC and DMC energies improved by 51 and 6 mEh respectively. It is exciting that our VMC optimization improves the orbitals even generated from LDA+U. But the most significant result is that with orbital optimization both wave function minimized to the same energy (similar to our FeO example in the previous section summarized in figure 4.3). Considering the decades of work and insight it took to produce qualitatively correct orbitals on the DFT front it is a significant result that our QMC calculation can make improvements.

Although we have demonstrated optimizing the orbitals within QMC will remove the starting point dependence of the orbitals and produce the best possible orbitals in the presence of the full interacting Hamiltonian for the minimal unit cell, for our more expensive supercell calculations it would be advantageous to have the best possible description of the orbitals to begin with. Work from others in the community indicate that HSE provides a good starting point choice for orbitals. [68, 97, 162] This functional incorporates exchange so that proper splitting of the d-orbitals occur, and self interaction is mitigated for both occupied and virtual orbitals. But it also produces an estimated band gap of 2.5 eV [162] which is larger than the gap produced by LDA+U (with reasonable U/J values). [100] Fur-

thermore G_0W_0 calculations of the band diagram using these orbitals show that there is only a slight shift of the bands compared to the DFT band diagram. [162] The minor shift of the band diagram suggests that the nonlocal screened-exchange contribution of HSE is similar to the G_0W_0 self-energy, and therefore HSE is a reliable functional.

4.6 MnO ($\Gamma \rightarrow \Gamma$)

With our wave function construction constraints chosen we perform our calculation for the $\Gamma \rightarrow \Gamma$ transition on a 16 atom super cell (which is a 2x2x2 super cell of the minimal 2 atom *chemical* unit cell (Figure 4.4).

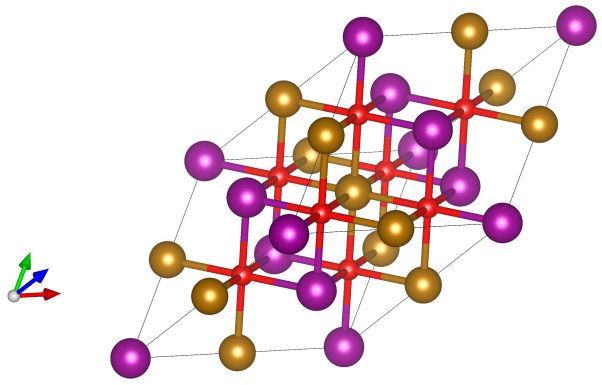


Figure 4.4: The 16 atom super cell (with atoms at the boundary displayed) illustrating the AF2 phase (an antiferromagnetic phase with planes of spin-up and spin-down polarized Mn atoms that alternate in the [111] direction). The purple and gold atoms correspond to spin-up and spin-dn polarized Mn atoms. The red atoms correspond to Oxygen atoms.

QMC MnO optical gap prediction

We compare our variance matched VMC-CIS-OO and DMC results to other current benchmarks in table 4.4. We found that our variance matched VMC-CIS-OO band gap overestimates the experimental gap and the subsequent DMC band gap result from the same wave functions lowers the gap by 1.6 eV. We suggest two errors that could contribute to the overestimate of the optical band gap prediction: (1) the value of the energy variance used to variance match is too high, and (2) the finite size correction for this particular system is inappropriate. In regard to the first error, if the energy variance for variance matching it too high it is less likely that a cancellation of error will occur between the ground and excited state. Therefore this band gap prediction at the very least suggests that correlation effects in the ground and excited state are different enough that they do not cancel entirely. In the

supplemental information it can be seen that the energy variance value used for matching is approximately ten times larger for MnO than that for MgO and PAE, but it is the best that could be computationally afforded. With respect to the second error, as mentioned in section 4.1 we addressed finite size errors for MnO in a manner different than that of MgO and PAE. For MnO we addressed finite size errors of the gap by performing simple single-particle hole gap DMC calculations with supercell sizes of 16, 32, and 64 atoms. Then we determined the correction to the 16 atom case that would be necessary to produce the extrapolated gap value, and applied this same correction to our VMC-CIS-OO gap and its corresponding DMC gap. The wave functions corresponding to the simple single-particle hole DMC calculations are not as sophisticated as that used for our VMC-CIS-OO calculation. With no orbital optimization and only a single-determinant the nodal surfaces of those calculations are certainly different than that of our VMC-CIS-OO wave functions. Therefore the slope of the extrapolation could be inappropriate to determine a finite size correction.

Method	gap (eV)
PBE	1.55
PBE+G_0W_0	1.39
HSE06	3.97
VMC-CIS-OO	6.5(3)
DMC	4.9(3)
Expt (opt. abs.)[144]	4.44

Table 4.4: The optical band gap of MnO determined by different methods.

QMC nature of MnO optical gap

Although we did not predict the experimental band gap of MnO to high accuracy we can provide insight into the nature of the optical excitation. Because CIS is the starting point for many excited state methods it can be argued that our VMC-CIS wave function captures the main qualitative physics of the exciton associated with the optical gap of MnO.

As stated earlier we have an active space for our VMC-CIS that includes all configurations that corresponds to the four types of transitions dissuced in literature. As we are working in the full interacting Hamiltoninan we can determine which of these transition correspond to the optical gap of MnO by analyzing the resultant VMC-CIS eigenvectors for the ground and first excited state.

The ground state had a weight of 94% for the reference determinant which suggests that HSE06 orbitals do qualitatively describe the ground state well. The excited state eigenvector had 14 determinants make up 93% of the wave function. Of those 14 determinants all corresponded (Mn 3d / O 2p) $\rightarrow$ (Mn 4s / O 3s). The analysis so far from the CI eigenvectors

suggest the transition to be of a mixture of (TM d → TM 4s) and (anion p → anion s). The second of these transitions is not one the four commonly suggested transitions listed earlier, but was discovered by Wagner and coworkers in 2004. [111]

It is well known that the 4s-orbital lies close in energy to the 3d-orbitals of isolated transition metal atoms. When considering the bulk MnO system ground state we can interpret the electrons of the 4s-orbital to be donated to the O 2p-orbital which creates the ionic ground state picture of Mn^{+2} O^{-2}. The 5 d-electrons take on their high spin configuration and cause a significant exchange splitting between the occupied and virtual d-orbitals. This pushes the occupied orbitals lower in energy and raises the virtual d-orbitals (both relative to the 4s-orbital) which is summarized pictorially in figure 4.5. Then when considering optical absorption we see that the transition is Mn 3d → Mn 4s. For simplicity we have ignored hybridization in this description of the excitation but as we describe next these transitions in reality involve hybridized states.

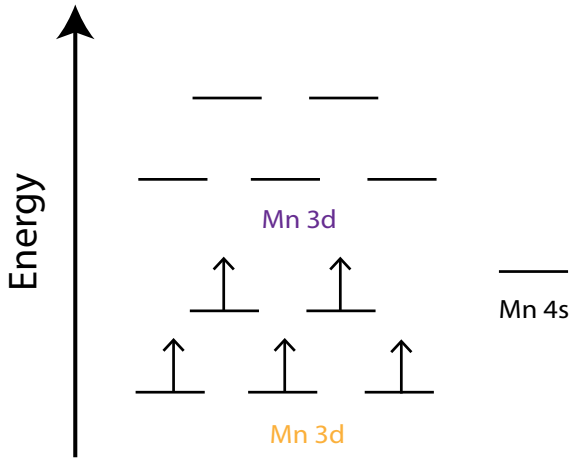


Figure 4.5: A sketch illustrating the high spin arrangement of d-electrons in d-orbitals determined by crystal field splitting of an octahedral geometry and the AF2 phase of MnO. The Mn 3d-orbitals are divided into two neighboring Mn sites (purple or gold) which reside in different ferromagnetic planes (refer to figure 4.4). The relative energy position of d-orbitals on neighboring Mn site is higher than that of the 4s Mn orbital.

To further characterise the exciton (correlated electron hole pair) associated with optical gap transition we use the single particle transition density matrix (T) defined by Martin as equation 4.1. [118] The matrix T is a N_o by N_v matrix in which N_o is the number of occupied orbitals, N_v is the number of virtual orbitals, Ψ_{ex} is the excited state wave function and Ψ_0 is the ground state wave function. The singular value decomposition of the matrix T results in a new set orbitals constructed via the unitary transformation determined by matrices U and V. This allows the T matrix to be expressed as equation 4.2 and the new set of orbitals are referred to as the Natural Transition Orbitals (NTO).

$$T_{ia} = \sum_{\sigma} < \Psi_{ex} | c^{\dagger}_{i\sigma} c_{a\sigma} | \Psi_0 > \tag{4.1}$$

$$[U^{\dagger} T V]_{ij} = \sqrt{\lambda_i} \delta_{ij} \tag{4.2}$$

The NTO decomposition provides a way of measuring the multi-configuartion nature of the excited state. In the case that a transition can be described simply between two orbitals the NTO analysis would result in a single singular value of $\lambda_{11} = 1$ and the picture of the particle and hole can be simply summarized as a single particle hole transition.

Our NTO analsysis of MnO results in this simple single particle hole transition which is pictorially summarized in figure 4.6. We refer to the particle-NTO as Ψ^e and the hole-NTO as Ψ^h. Notice that the hole takes on a form that one would expect from a d^2sp^3 hybridization of orbitals on a Mn site in an octahedral arrangement, mixed with some 2p character from the surrounding oxygens. There is symmetry in how the lobes of the d^2sp^3 point towards the 2p oxygen and this motif repeats throughout the ferromagnetic plane and between all ferromagnetic planes uniformly. Just like our CI eigenvector analysis found the NTO depicts the optical transition to be a mixture of (TM d $\rightarrow$ TM 4s) and (anion p $\rightarrow$ anion s).

One may wonder "if the MnO optical gap can be so simply summarized is there a functional that can capture this?". To answer this question we compared the particle hole pairs produced from the most commonly used functionals (which include LDA, LDA+U, HF, HSE06) and found that while they all surprisingly produce consist description of Ψ^e with each other and with our NTO-Ψ^e none produced a Ψ^h consistent with our NTO-Ψ^h. During the comparison of these functionals we found that the HOMO (ψ_H), HOMO-1 (ψ_{H-1}), HOMO-2 (ψ_{H-2}) orbitals were degenerate (explicitly so for LDA+U and HSE06 while there was a minor difference of 10^{-4} eV for HF and LDA) and so there was a question of whether a simple mixing of these degenerate orbitals could reproduce the NTO-Ψ^e/NTO-Ψ^h. To answer that question we performed a least squares fit to equation 4.3 to investigate how well the NTO hole could be described by the highest occupied orbitals produced by these functionals.

$$||(a * \psi^2_H + b * \psi^2_{H-1} + c * \psi^2_{H-2}) - (\Psi^h)^2||^2 \tag{4.3}$$

The resultant holes from minimizing equation 4.3 for LDA, LDA+U, HSE06, and HF are summarized in figure 4.7 with HSE06 providing the best fit. All functionals seem to be able to reproduce the shape of the d-orbital on the center Mn from our NTO-Ψ^h but none reproduce the shape of the orbitals centered on the oxygen atoms from our NTO-Ψ^h. The hole orbital produced by the set of functionals commonly have sp characer on the oxygen ligands. Note that for these sp orbitals the lobe closer to the Mn site is larger than the one further and that the difference in their sizes decreases in the order of LDA, LDA+U, HSE06 and HF. At the extreme of HF the oxygen orbitals seem more like 2p character than sp character. Another trend moving from LDA through to HF is that as the 2p character of

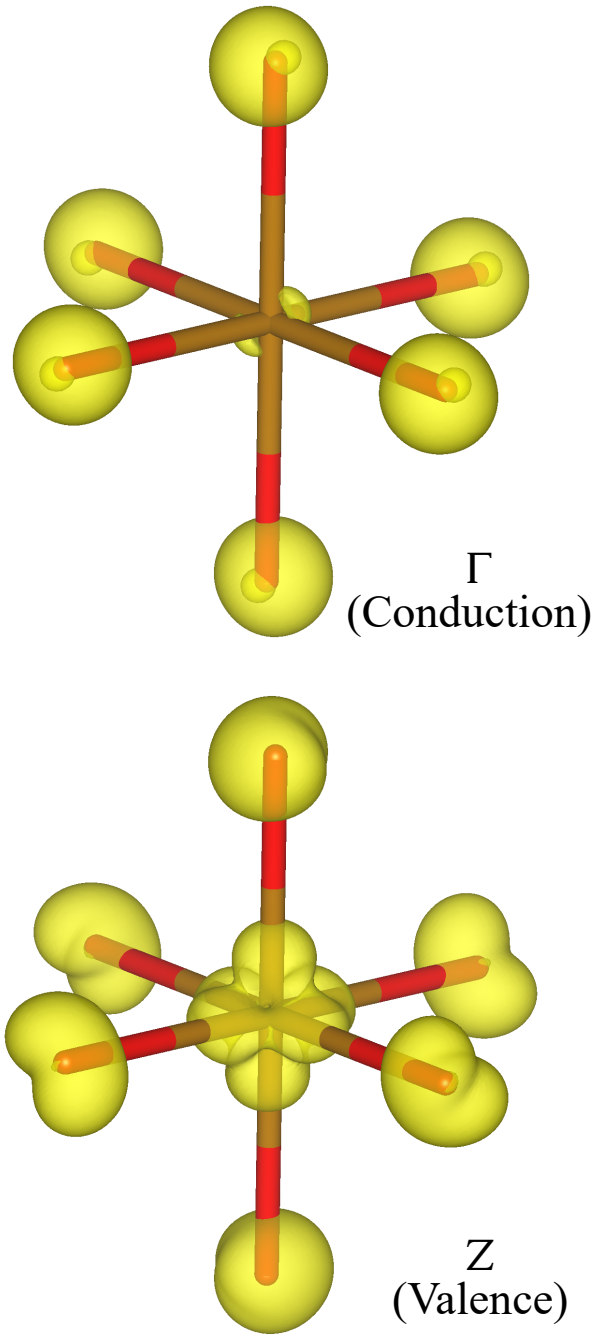


Figure 4.6: The NTO-VMC isosurfaces associated with the particle (Ψ^e) in the conduction band and the hole (Ψ^h) in the valence band.

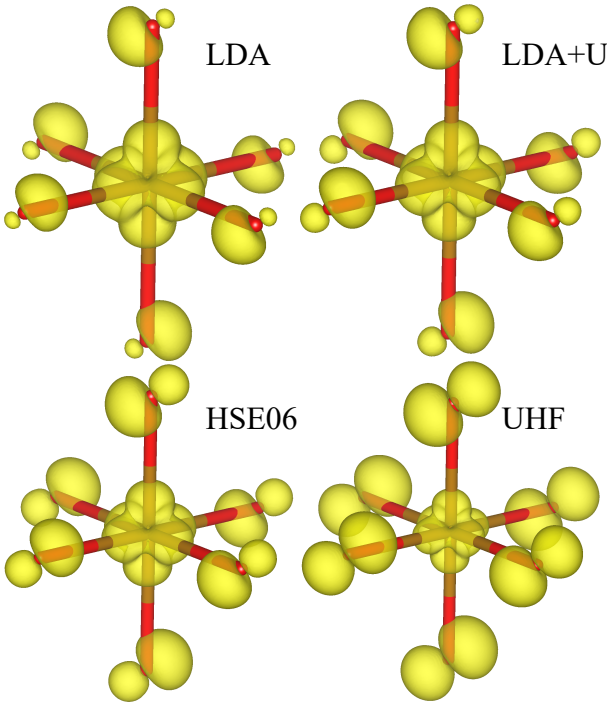


Figure 4.7: Comparison of the hole (Ψ^h) in the valence band from various single particle theories.

the oxygen orbitals grows the d amplitude of the center Mn orbitals seems to be shrinking. When compared to NTO-Ψ^h LDA and LDA+U put too much density on the center Mn while UHF puts to little and HSE06 has an intermediate amount most similar to the NTO-Ψ^h. These trends are consistent with the findings of Wagner who noted that the modulation of exact exchange mixing causes a modulation of the hybridization of the orbitals.[111] So it seems that the VMC-CIS-OO wave function can correct the degree of hybridization but also fundamentally changes the shape of the orbitals centered on the oxygens to some degree.

4.7 MnO ($Z \rightarrow \Gamma$)

For the $Z \rightarrow \Gamma$ transition we performed calculation on the 32 atom supercell (which is the 2x2x2 supercell of the minimal 4 atom *magnetic* unit cell). Note that this supercell is the minimal unit cell that can capture the indirect band gap transition.

QMC MnO optical gap prediction

The 32 atom supercell presents a difficult system to perform computations on, limiting the techniques we can apply to it for study. We produce a prediction for the gap using variance matched VMC-CIS(D) wave function. Then we evaluate various SPO sets for use in G_0W_0 calculations. The results of our VMC and G_0W_0 calculations are in table 4.5.

For a VMC-CIS(D)-OO calculation the number of parameters necessary for orbital optimization alone requires more RAM than current standard HPC resources offer. Therefore we cannot perform orbital optimization for this system and instead only perform a VMC-CIS(D) calculation. But we can still have a unique set of orbitals for the ground state and excited state by simply using orbitals from different functionals. So as a substitution to explicit rotation of orbital on the VMC side we vary the fraction of exact exchange for hybrid functionals when producing SPO sets and evaluate the best exact exchange fraction to use for the ground state and the excited state individually using VMC and variance matching. As we noted in our study of the $\Gamma \rightarrow \Gamma$ transition the HSE06 orbitals provided a good description of the ground state, so we decided to use HSE06 orbitals with standard screening and exchange parameters ($\omega = 0.20$ and a $= 0.25$) for the ground state. To variance match the ground state we varied the amount of exact exchange involved in the excited state and found that HSE06 parameters of ($\omega = 0.20$ and a $= 0.35$) did well to explicitly variance match (refer to appendix B.1).

We evaluate the quality of various SPO sets for a G_0W_0 calculation in the same manner as was done previously by Zhao and Neuscamman.[228] The steps are briefly summarized as: 1) select a functional, 2) produce a histogram of $\Delta_{ia}^{DFT} - \Delta_{ia}^{VMC}$ (in which Δ_{ia}^{DFT} is the gap predicted by DFT between Kohn-Sham states indexed by 'i' and 'a' and Δ_{ia}^{VMC} is the gap predicted by VMC using SJ wave functions for the ground/excited states with a simple single promotion from orbital 'a' to orbital 'i' in the excited state) 3) repeat steps 1/2 with another functional until a histogram centered around zero with small variance is found. The motivation for this method of selecting the best functional is that quasiparticle energies are computed according to equation 1.22, so 'more realistic' Kohn-Sham eigenvalues would require less correction. Because VMC gaps are computed using the full interacting Hamiltonian we use those gaps as a the standard for 'correct' when evaluating Kohn-Sham eigenvalues. Although in reality the excitations from VMC correspond to optical absorption energies while the quasiparticle energies correspond to photo emission energies so it is not quite a fair comparison. To construct the histograms for MnO we simply use a PBE0 functional while varying the fraction of exact exchange (examples of these histograms can be found in appendix B.1). We found that an exchange fraction of 0.55 produces a histogram centered around zero, and a choice of 0.45 produced a histogram skewed the right of zero. But because the transition of interest $\Delta_{LUMO,HOMO}^{DFT} - \Delta_{LUMO,HOMO}^{VMC}$ was smallest for an exchange fraction of 0.45 this exchange fraction is chosen to produce the SPO set for G_0W_0.

Just as it was for the $\Gamma \rightarrow \Gamma$ transition, our VMC-CISD gap estimation is an overestimate of the $Z \rightarrow \Gamma$ transition gap. We explain the overestimation with the same shortcomings that the $\Gamma \rightarrow \Gamma$ study revealed. But in the case of $Z \rightarrow \Gamma$ the short comings are more severe.

Method	gap (eV)
UHF [209]	12.9
GGA [162]	0.7
GGA+U [162]	1.3
HSE03 [162]	2.5
GGA+G_0W_0 [162]	1.6
HSE03+G_0W_0 [162]	3.2
HSE06	3.29
VMC	6.31
VMC-CISD	4.87
PBE(0.45)	6.33
PBE(0.45)+G_0W_0	4.90
Expt (photoem.) [53]	3.9(4)
Expt (cond.) [48]	3.8..4.2
Expt (opt. abs.) [96]	3.6..3.8
Expt (eng. loss.) [67]	3.1 (or 4.9)
Expt (opt. abs.) [144]	3.4

Table 4.5: The optical band gap of MnO determined by different methods.

Just as before, the variance used for variance matching could be much too high to have a cancellation of the correlation errors between the ground and excited state. In this case no orbital relaxation could occur so there was less flexibility in these wave functions than there was in the wave functions used to investigate $\Gamma \to \Gamma$.

QMC nature of MnO optical gap

The orbitals participating in the transition of $Z \to \Gamma$ are not qualitatively different from those involved in $\Gamma \to \Gamma$ so we forego the same analysis for this transition because the same conclusions will be drawn. Instead we present some thought provoking questions of why the gap of MnO is difficult to predict quantitatively.

Are the ground or lowest excited state of MnO strongly correlated?

Our VMC-CIS(D) calculation with HSE06 orbitals suggest that neither the ground nor excited state is strongly correlated. Performing VMC-CISD calculations does not contribute additional configurations with significant weight for the ground state or excited state. Both states are described well with a single Slater Jastrow wave functions. Of course there is an unknown factor of what would happen if the active space were increased and if excitation

of higher degree were included. But our VMC-CISD calculation utilizes a reasonable active space and we cannot provide a chemical intuitive explanation of any deficiency.

If both states are weakly correlated are we completely capturing that weak correlation?

The one-body and two-body Jastrow factors used for our calculation of MnO contain only a few terms of the expansion of the Jastrow factor in many body terms ($U_{ee} + U_{en} + U_{een} + ...$). Therefore we are certainly not capturing all the available dynamic correlation. In the limit of including all the many body terms we would simply have the DMC result and are limited by the quality of the nodal surface of the trail wave function.

If weak correlation is captured sufficiently why are the gap predictions overestimating?

Our scheme of variance matching the ground and excited state only works at the limit of the variance approaching zero. When the variance is far from zero there is no guarantee that the ground state and excited state each have captured an amount of correlation energy that makes the gap accurate. To approach the limit of zero variance we simply need to continue variationally improving the wave functions, but this of course leads to the need for an exponential amount of parameters.

Why not perform variance matching using DMC energy variances?

One might suggest performing DMC calculations to work with lower energy variances and capture significant additional weak correlation, but we argue that it is premature to pursue that avenue at this time.

Wagner and coworkers have already attempted to use DMC to predict the gap but ultimately produced an overestimation.[180] They noted that varying the fraction of exact-exchange varies the degree of hybriziation of orbitals significantly, which vary the nodal surface and therefore affect the DMC outcome.[180] But their technique of scanning values of exact exchange fraction to find the variationally lowest DMC energy for the ground state does not guarantee that the SPO set produces equal quality nodal surfaces for the ground and excited state.

One may hope that because we used VMC variance matching to determine separate values of exact exchange fraction for the ground and excited state SPO sets that the quality of the nodal surfaces of the ground and excited state will now be more equal, but this is far from being a guarantee. The first issue is that variance matching the VMC energies will not necessarily lead to variance matched DMC energies. At this point one may say to simply perform multiple DMC energy evaluations with various values of exact exchange fraction being used to perform variance matching. That is an incredibly expensive task to do for MnO and has not been proven to be fruitful with smaller systems yet. That approach also has an implication that scanning the exact exchange fraction is the same a scanning

all plausible SPO sets that can be used, which is not true. We mentioned in section 1.4 that hybrid functionals are not free of self-interaction error[8] and it is abundantly clear through our work (and the work of others[180]) that the localization of electron density (which self interaction error affects) is what needs to be address for accurate SPO description of ground and excited MnO states. Therefore simply scanning the exact exchange fraction for DMC variance matching in all liklihood could be insufficient to produce a quantitiatively correct optical gap. In addition performing DMC on multiple functionals with multiple parameterizations is a brute force 'needle in a haystack' approach to the band gap of MnO which is what our work endeavours to avoid.

Chapter 5

Conclusion

We have shown in this dissertation that orbital optimization of Multi-Slater Jastrow wave functions and its coupling to an excited state targeting function, configuration selection for QMC wave functions, a variance matching scheme for optical gaps, and a modified guiding function for sampling within QMC produce reliable optical gaps.

For aperiodic systems we found that our QMC optical gap workflow produces predictions on par in terms of accuracy with other standard techniques (e.g. MRCI+Q, CASSCF, CASPT2, EOM-CCSD) for small molecules (e.g. Formaldimine, Thioformaldehyde). The modified guiding function was found to be quite useful and necessary to produce energy variance values with low standard errors, which improved both the optimization and the reliability of our variance matching scheme. Furthermore, we found that the configurations captured through S-CI contained many that were not capture through a standard active space choice. Finally, when constructing VMC wave functions for the ground and excited state, simply using a threshold weight to select configurations from S-CI calculations to construct the wave function did not produce vertical excitations values as accurate as the implicit variance matching scheme.

The case of $\left[C_3N_2O_2H_4Cl\right]^-$ is a larger stress for standard techniques and provided opportunity for our workflow to demonstrate its advantages over active space methods for various reasons. It's size made it so that brute force large active space would not be enough to capture dynamic correlation and the large dipole change from ground to excited state would mean a state averaging of the orbitals most likely would not allow for appropriate relaxation for either the ground or excited state. During the CASSCF+CASPT2 workflow we found that CASSCF with various active space choice, number of states and weighting of these state all created a zeroth order starting point that produced 'intruder states' during the CASPT2 calculation. Therefore a choice in shift value and type resulted in a range of optical gaps from 2.127 eV to 4.669 eV being produced. Meanwhile, our explicit variance matched VMC calculation produced an estimate of the vertical excitation on par with EOM-CCSD when 200 or 1000 configurations were included for the excited state.

For periodic systems we found that our QMC optical gap workflow produces predictions on par in terms of accuracy with other standard techniques (e.g. DFT, G_0W_0) for simple

bulk materials (e.g. MgO, Trans-Polyaceylene). Bulk MgO was a system that DFT could predict the optical gap for easily. With a great SPO set from DFT, G_0W_0 and our approach could also produce estimates of a gap similar to the experimentally reported value. The more challenging Trans-Polyaceylene proves difficult for hybird functional DFT to get right. The HSE06 estimate of a gap of 0.53 eV suggests that the functional has shortcomings, but both G_0W_0 and VMC-CIS(D)-OO can correct the gap to be on par with experimentally reported value. For bulk FeO we show that VMC orbital optimization reduced the starting point dependence of the SPO set for DMC. For bulk MnO we produce an overestimate of the optical gap and are no closer in accuracy than DFT, G_0W_0 or past QMC work. But we performed an in depth analysis of the nature of the exciton for the opitical gap which allows us to state that the lowest transition is a mixture of (TM d $\rightarrow$ TM 4s) and (anion p $\rightarrow$ anion s). Our close analysis also explains that this system is so difficult for a wide variety of methods due to the difficulty of capturing its weak correlation.

The work in this dissertation suggests that future directions could include improving the scaling of optimization of VMC parameters, or creating a S-CI native to VMC to take into account the affect of the Jastrow factor in describing weak correlation. Both directions would help improve the ability to perform VMC on large systems for accurate optical gaps. The first direction would help allow orbital optimization on larger systems. As stated in section 4.7 the amount of parameters necessary to perform orbital optimization for systems like the 32 atom MnO supercell simply exceed what current computational resources can handel. The second direction, a VMC native S-CI, would eliminate the need for QMC to interface with other active space methods, and be more efficient than VMC-CIS(D). An active space method with orbital optimization native to QMC would make a truly standalone QMC method (QMC-CASSCF).

Bibliography

[1] A. Scemama et al. "Quantum Package v1.1". In: (). DOI: 10.5281/zenodo.825872. DOI: 10.5281/zenodo.825872.

[2] ABINIT. *The generation of atomic data.* [Online; accessed April 24, 2019]. 2019. URL: https://www.abinit.org/sites/default/files/infos/8.6/tutorial/generated_files/lesson_paw2.html.

[3] A.A. Abrikosov, L.P. Gorkov, and E. Dzyaloshinskii. *Methods of quantum field theory in statistical physics*. Dover, New-York, 1975.

[4] Paulo H. Acioli and David M. Ceperley. "Generation of pseudopotentials from correlated wave functions". In: *The Journal of Chemical Physics* 100.11 (1994), pp. 8169–8177. DOI: 10.1063/1.466811. eprint: https://doi.org/10.1063/1.466811. URL: https://doi.org/10.1063/1.466811.

[5] Carlo Adamo and Vincenzo Barone. "Exchange functionals with improved long-range behavior and adiabatic connection methods without adjustable parameters: The m and m models". In: *The Journal of Chemical Physics* 108.2 (Jan. 1998), pp. 664–675. ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.475428. URL: http://dx.doi.org/10.1063/1.475428.

[6] Reinhart Ahlrichs. "Methods for efficient evaluation of integrals for Gaussian type basis sets". In: *Theoretica chimica acta* 33.2 (1974), pp. 157–167.

[7] H Ahmad et al. "Hydrogen from photo-catalytic water splitting process: A review". In: *Renewable and Sustainable Energy Reviews* 43 (2015), pp. 599–610.

[8] C-O Almbladh and Ulf von Barth. "Exact results for the charge and spin densities, exchange-correlation potentials, and density-functional eigenvalues". In: *Physical Review B* 31.6 (1985), p. 3231.

[9] James B Anderson. "A random-walk simulation of the Schrödinger equation: H+ 3". In: *The Journal of Chemical Physics* 63.4 (1975), pp. 1499–1503.

[10] James B Anderson. "Quantum chemistry by random walk." In: *The Journal of Chemical Physics* 65.10 (1976), pp. 4121–4127.

[11] Vladimir I Anisimov, Ferdi Aryasetiawan, and AI Lichtenstein. "First-principles calculations of the electronic structure and spectra of strongly correlated systems: the LDA+ U method". In: *Journal of Physics: Condensed Matter* 9.4 (1997), p. 767.

[12] Roland Assaraf, S. Moroni, and Claudia Filippi. "Optimizing the Energy with Quantum Monte Carlo: A Lower Numerical Scaling for Jastrow–Slater Expansions". In: *J. Chem. Theory Comput.* 13.11 (2017), pp. 5273–5281. DOI: 10.1021/acs.jctc.7b00648.

[13] Rodney J Bartlett and Isaiah Shavitt. "Determination of the size-consistency error in the single and double excitation configuration interaction model". In: *International Journal of Quantum Chemistry* 12.S11 (1977), pp. 165–173.

[14] Federico Becca and Sandro Sorella. *Quantum Monte Carlo approaches for correlated systems*. Cambridge University Press, 2017.

[15] Axel D. Becke. "A new mixing of Hartree Fock and local density functional theories". In: *The Journal of Chemical Physics* 98.2 (1993), pp. 1372–1377. DOI: 10.1063/1.464304. eprint: https://doi.org/10.1063/1.464304. URL: https://doi.org/10.1063/1.464304.

[16] Lorin X. Benedict, Eric L. Shirley, and Robert B. Bohn. "Optical Absorption of Insulators and the Electron-Hole Interaction: An Ab Initio Calculation". In: *Phys. Rev. Lett.* 80 (20 May 1998), pp. 4514–4517. DOI: 10.1103/PhysRevLett.80.4514. URL: https://link.aps.org/doi/10.1103/PhysRevLett.80.4514.

[17] M. Chandler Bennett et al. "A new generation of effective core potentials from correlated calculations: 2nd row elements". In: *The Journal of Chemical Physics* 149.10 (2018), p. 104108. DOI: 10.1063/1.5038135. eprint: https://doi.org/10.1063/1.5038135. URL: https://doi.org/10.1063/1.5038135.

[18] Peter E. Blöchl. "Generalized separable potentials for electronic-structure calculations". In: *Phys. Rev. B* 41 (8 Mar. 1990), pp. 5414–5416. DOI: 10.1103/PhysRevB.41.5414. URL: https://link.aps.org/doi/10.1103/PhysRevB.41.5414.

[19] Nick S Blunt and Eric Neuscamman. "Charge-transfer excited states: Seeking a balanced and efficient wave function ansatz in variational Monte Carlo". In: *J. Chem. Phys.* 147.19 (2017), p. 194101.

[20] Nick S Blunt and Eric Neuscamman. "Excited-state diffusion Monte Carlo calculations: a simple and efficient two-determinant ansatz". In: *arXiv* 1808.09549 (2018).

[21] J H de Boer and E J W Verwey. "Semi-conductors with partially and with completely filled3d-lattice bands". In: *Proceedings of the Physical Society* 49.4S (Aug. 1937), pp. 59–71. DOI: 10.1088/0959-5309/49/4s/307. URL: https://doi.org/10.1088%2F0959-5309%2F49%2F4s%2F307.

[22] Max Born, Werner Heisenberg, and Pascual Jordan. "Zur Quantenmechanik. II." In: *Zeitschrift für Physik* 35.8-9 (1926), pp. 557–615.

[23] Max Born and Pascual Jordan. "Zur quantenmechanik". In: *Zeitschrift für Physik* 34.1 (1925), pp. 858–888.

[24] Max Born and Robert Oppenheimer. "Zur quantentheorie der molekeln". In: *Annalen der physik* 389.20 (1927), pp. 457–484.

[25] B.H. Brandow. "Electronic structure of Mott insulators". In: *Advances in Physics* 26.5 (1977), pp. 651–808. DOI: 10.1080/00018737700101443. eprint: https://doi.org/10.1080/00018737700101443. URL: https://doi.org/10.1080/00018737700101443.

[26] Thomas Bredow and Andrea R. Gerson. "Effect of exchange and correlation on bulk properties of MgO, NiO, and CoO". In: *Phys. Rev. B* 61 (8 Feb. 2000), pp. 5194–5201. DOI: 10.1103/PhysRevB.61.5194. URL: https://link.aps.org/doi/10.1103/PhysRevB.61.5194.

[27] Fabien Bruneval and Xavier Gonze. "Accurate GW self-energies in a plane-wave basis using only a few empty states: Towards large systems". In: *Phys. Rev. B* 78.8 (Aug. 2008), p. 085125. ISSN: 1098-0121, 1550-235X. DOI: 10.1103/physrevb.78.085125. URL: https://doi.org/10.1103/physrevb.78.085125.

[28] Fabien Bruneval, Nathalie Vast, and Lucia Reining. "Effect of self-consistency on quasiparticles in solids". In: *Phys. Rev. B* 74.4 (June 2006), p. 045102. ISSN: 1098-0121, 1550-235X. DOI: 10.1103/physrevb.74.045102. URL: https://doi.org/10.1103/physrevb.74.045102.

[29] M. Burkatzki, C. Filippi, and M. Dolg. "Energy-consistent pseudopotentials for quantum Monte Carlo calculations". In: *J. Chem. Phys.* 126.23 (2007), p. 234105. DOI: 10.1063/1.2741534.

[30] M Burkatzki, C Filippi, and M Dolg. "Energy-consistent pseudopotentials for quantum Monte Carlo calculations". In: *The Journal of chemical physics* 126.23 (2007), p. 234105.

[31] J. H. Burroughes et al. "Light-emitting diodes based on conjugated polymers". In: *Nature* 347.6293 (1990), pp. 539–541. ISSN: 1476-4687. DOI: 10.1038/347539a0. URL: https://doi.org/10.1038/347539a0.

[32] Elena Bykova et al. "Structural complexity of simple Fe 2 O 3 at high pressures and temperatures". In: *Nature communications* 7.1 (2016), pp. 1–6.

[33] Laimutis Bytautas and Klaus Ruedenberg. "A priori identification of configurational deadwood". In: *Chemical Physics* 356.1-3 (2009), pp. 64–75.

[34] Damian Carrieri et al. "Photo-catalytic conversion of carbon dioxide to organic acids by a recombinant cyanobacterium incapable of glycogen storage". In: *Energy & Environmental Science* 5.11 (2012), pp. 9457–9461.

[35] C. Castiglioni, G. Zerbi, and M. Gussoni. "Peierls distortion in trans polyacetylene: Evidence from infrared intensities". In: *Solid State Communications* 56.10 (1985), pp. 863–866. ISSN: 0038-1098. DOI: https://doi.org/10.1016/0038-1098(85)90421-1. URL: http://www.sciencedirect.com/science/article/pii/0038109885904211.

[36] Michele Casula, Claudio Attaccalite, and Sandro Sorella. "Correlated geminal wave function for molecules: An efficient resonating valence bond approach". In: *The Journal of chemical physics* 121.15 (2004), pp. 7110–7126.

[37] D. M. Ceperley and B. J. Alder. "Ground State of the Electron Gas by a Stochastic Method". In: *Phys. Rev. Lett.* 45 (7 Aug. 1980), pp. 566–569. DOI: 10.1103/PhysRevLett.45.566. URL: https://link.aps.org/doi/10.1103/PhysRevLett.45.566.

[38] David Ceperley, Geoffrey V Chester, and Malvin H Kalos. "Monte Carlo simulation of a many-fermion study". In: *Physical Review B* 16.7 (1977), p. 3081.

[39] M. Chandross, Y. Shimoi, and S. Mazumdar. "Diagrammatic exciton-basis valence-bond theory of linear polyenes". In: *Phys. Rev. B* 59 (7 Feb. 1999), pp. 4822–4838. DOI: 10.1103/PhysRevB.59.4822. URL: https://link.aps.org/doi/10.1103/PhysRevB.59.4822.

[40] Shu-Wei Chang and Henryk A Witek. "Choice of optimal shift parameter for the intruder state removal techniques in multireference perturbation theory". In: *Journal of chemical theory and computation* 8.11 (2012), pp. 4053–4061.

[41] Jong H Choi, Charles F Lebeda, and Richard P Messmer. "Variational principle for excited states: Exact formulation and other extensions". In: *Chem. Phys. Lett.* 5.8 (1970), pp. 503–506.

[42] Bryan K. Clark et al. "Computing the energy of a water molecule using multideterminants: A simple, efficient algorithm". In: *J. Chem. Phys.* 135.24 (2011), p. 244105. DOI: 10.1063/1.3665391.

[43] D J Clouthier and D A Ramsay. "The Spectroscopy of Formaldehyde and Thioformaldehyde". In: *Annu. Rev. Phys. Chem.* 34.1 (1983), pp. 31–58. DOI: 10.1146/annurev.pc.34.100183.000335.

[44] Matteo Cococcioni and Stefano de Gironcoli. "Linear response approach to the calculation of the effective interaction parameters in the LDA+U method". In: *Phys. Rev. B* 71 (3 Jan. 2005), p. 035105. DOI: 10.1103/PhysRevB.71.035105. URL: https://link.aps.org/doi/10.1103/PhysRevB.71.035105.

[45] Monika Dash et al. "Perturbatively selected configuration-interaction wave functions for efficient geometry optimization in quantum Monte Carlo". In: *J. Chem. Theory Comput.* 14 (2018), p. 4176. DOI: 10.1021/acs.jctc.8b00393.

[46] P. A. M. Dirac. "Quantum Mechanics of Many-Electron Systems". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 123.792 (1929), pp. 714–733. ISSN: 09501207. URL: http://www.jstor.org/stable/95222.

[47] Paul Adrien Maurice Dirac. "The quantum theory of the electron". In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 117.778 (1928), pp. 610–624.

[48] I A Drabkin et al. In: *Fiz. Tverd. Tela (Leningrad)* 10 (1968), p. 3082.

[49] N. D. Drummond, M. D. Towler, and R. J. Needs. "Jastrow correlation factor for atoms, molecules, and solids". In: *Phys. Rev. B* 70 (23 Dec. 2004), p. 235119. DOI: 10.1103/PhysRevB.70.235119. URL: https://link.aps.org/doi/10.1103/PhysRevB.70.235119.

[50] N. D. Drummond, J. R. Trail, and R. J. Needs. "Trail-Needs pseudopotentials in quantum Monte Carlo calculations with plane-wave/blip basis sets". In: *Phys. Rev. B* 94 (16 Oct. 2016), p. 165170. DOI: 10.1103/PhysRevB.94.165170. URL: https://link.aps.org/doi/10.1103/PhysRevB.94.165170.

[51] Neil David Drummond, JR Trail, and RJ Needs. "Trail-Needs pseudopotentials in quantum Monte Carlo calculations with plane-wave blip basis sets". In: *Physical Review B* 94.16 (2016), p. 165170.

[52] Christophe Dugave. *Cis-trans isomerization in biochemistry*. Weinheim: Wiley-VCH, 2006. ISBN: 9783527609338.

[53] J. van Elp et al. "Electronic structure of MnO". In: *Phys. Rev. B* 44 (4 July 1991), pp. 1530–1537. DOI: 10.1103/PhysRevB.44.1530. URL: https://link.aps.org/doi/10.1103/PhysRevB.44.1530.

[54] S. Fahy, X. W. Wang, and Steven G. Louie. "Variational Quantum Monte Carlo Nonlocal Pseudopotential Approach to Solids: Cohesive and Structural Properties of Diamond". In: *Phys. Rev. Lett.* 61 (14 Oct. 1988), pp. 1631–1634. DOI: 10.1103/PhysRevLett.61.1631. URL: https://link.aps.org/doi/10.1103/PhysRevLett.61.1631.

[55] S. Fahy, X. W. Wang, and Steven G. Louie. "Variational quantum Monte Carlo nonlocal pseudopotential approach to solids: Formulation and application to diamond, graphite, and silicon". In: *Phys. Rev. B* 42 (6 Aug. 1990), pp. 3503–3522. DOI: 10.1103/PhysRevB.42.3503. URL: https://link.aps.org/doi/10.1103/PhysRevB.42.3503.

[56] B Scott Fales and Todd J Martinez. "Efficient treatment of large active spaces through multi-GPU parallel implementation of direct configuration interaction". In: *Journal of Chemical Theory and Computation* 16.3 (2020), pp. 1586–1596.

[57] Ugo Fano. "Effects of configuration interaction on intensities and phase shifts". In: *Physical Review* 124.6 (1961), p. 1866.

[58] Willliam Feller. *An introduction to probability theory and its applications, vol 2*. John Wiley & Sons, 2008.

[59] JJ Fernandez. "G W calculations in an exactly solvable model system at different dilution regimes: The problem of the self-interaction in the correlation part". In: *Physical Review A* 79.5 (2009), p. 052513.

[60] A.L. Fetter and J.D. Walecka. *Quantum Theory of Many-Particle Systems*. McGraw-Hill, New York, 1971.

[61] RP Feynman and Michael Cohen. "Energy spectrum of the excitations in liquid helium". In: *Physical Review* 102.5 (1956), p. 1189.

[62] Claudia Filippi, Roland Assaraf, and Saverio Moroni. "Simple formalism for efficient derivatives and multi-determinant expansions in quantum Monte Carlo". In: *J. Chem. Phys.* 144.19 (2016), p. 194105. DOI: 10.1063/1.4948778.

[63] W. M. C. Foulkes et al. "Quantum Monte Carlo simulations of solids". In: *Rev. Mod. Phys.* 73 (1 Jan. 2001), pp. 33–83. DOI: 10.1103/RevModPhys.73.33. URL: https://link.aps.org/doi/10.1103/RevModPhys.73.33.

[64] WMC Foulkes et al. "Quantum Monte Carlo simulations of solids". In: *Rev. Mod. Phys.* 73.1 (2001), p. 33.

[65] Irmgard Frank et al. "Molecular dynamics in low-spin excited states". In: *J. Chem. Phys.* 108.10 (1998), pp. 4060–4069. DOI: 10.1063/1.475804.

[66] R. H. Friend et al. "Electroluminescence in conjugated polymers". In: *Nature* 397.6715 (1999), pp. 121–128. ISSN: 1476-4687. DOI: 10.1038/16393. URL: https://doi.org/10.1038/16393.

[67] Bärbel Fromme, U Brunokowski, and E Kisker. "d- d excitations and interband transitions in MnO: A spin-polarized electron-energy-loss study". In: *Physical Review B* 58.15 (1998), p. 9783.

[68] Alejandro J. Garza and Gustavo E. Scuseria. "Predicting Band Gaps with Hybrid Density Functionals". In: *The Journal of Physical Chemistry Letters* 7.20 (2016), pp. 4165–4170. DOI: 10.1021/acs.jpclett.6b01807. eprint: https://doi.org/10.1021/acs.jpclett.6b01807. URL: https://doi.org/10.1021/acs.jpclett.6b01807.

[69] Giovanni Ghigo, Björn O. Roos, and Per-Åke Malmqvist. "A modified definition of the zeroth-order Hamiltonian in multiconfigurational perturbation theory (CASPT2)". In: *Chem. Phys. Lett.* 396.1 (2004), pp. 142–149. ISSN: 0009-2614. DOI: https://doi.org/10.1016/j.cplett.2004.08.032. URL: http://www.sciencedirect.com/science/article/pii/S0009261404012242.

[70] Paolo Giannozzi et al. "QUANTUM ESPRESSO: a modular and open-source software project for quantum simulations of materials". In: *Journal of physics: Condensed matter* 21.39 (2009), p. 395502.

[71] Emmanuel Giner, Anthony Scemama, and Michel Caffarel. "Using perturbatively selected configuration interaction in quantum Monte Carlo calculations". In: *Can. J. Chem.* 91.9 (2013), p. 879.

[72] S. Goedecker, M. Teter, and J. Hutter. "Separable dual-space Gaussian pseudopotentials". In: *Phys. Rev. B* 54.3 (July 1996), pp. 1703–1710. ISSN: 0163-1829, 1095-3795. DOI: 10.1103/physrevb.54.1703. URL: https://doi.org/10.1103/physrevb.54.1703.

[73] Xavier Gonze, Peter Käckell, and Matthias Scheffler. "Ghost states for separable, norm-conserving, Iab initioP pseudopotentials". In: *Phys. Rev. B* 41 (17 June 1990), pp. 12264–12267. DOI: 10.1103/PhysRevB.41.12264. URL: https://link.aps.org/doi/10.1103/PhysRevB.41.12264.

[74] Mark S Gordon and Michael W Schmidt. "Advances in electronic structure theory: GAMESS a decade later". In: *Theory and applications of computational chemistry*. Elsevier, 2005, pp. 1167–1189.

[75] Andreas Griewank and Andrea Walther. *Evaluating Derivatives, Principles and Techniques of Algorithmic Differentiation, Second Edition*. Philadelphia: Society for Industrial and Applied Mathematics, 2008.

[76] The Abinit group. *Many-Body Theory in ABINIT*. URL: https://docs.abinit.org/theory/mbt/.

[77] DR Hamann, M Schlüter, and C Chiang. "Norm-conserving pseudopotentials". In: *Physical Review Letters* 43.20 (1979), p. 1494.

[78] Brian L. Hammond, Peter J. Reynolds, and William A. Lester. "Valence quantum Monte Carlo with ab initio effective core potentials". In: *The Journal of Chemical Physics* 87.2 (1987), pp. 1130–1136. DOI: 10.1063/1.453345. eprint: https://doi.org/10.1063/1.453345. URL: https://doi.org/10.1063/1.453345.

[79] Brian L Hammond, William A Lester, and Peter James Reynolds. *Monte Carlo methods in ab initio quantum chemistry*. Vol. 1. World Scientific, 1994.

[80] Richard A Heaton, Joseph G Harrison, and Chun C Lin. "Self-interaction correction for density-functional theory of electronic energy bands of solids". In: *Physical Review B* 28.10 (1983), p. 5992.

[81] Lars Hedin. "New Method for Calculating the One-Particle Green's Function with Application to the Electron-Gas Problem". In: *Phys. Rev.* 139.3A (Aug. 1965), A796–A823. ISSN: 0031-899X. DOI: 10.1103/physrev.139.a796. URL: https://doi.org/10.1103/physrev.139.a796.

[82] Warren J Hehre, Robert F Stewart, and John A Pople. "self-consistent molecular-orbital methods. i. use of gaussian expansions of Slater-type atomic orbitals". In: *The Journal of Chemical Physics* 51.6 (1969), pp. 2657–2664.

[83] Walter Heitler and Fritz London. "Wechselwirkung neutraler Atome und homöopolare Bindung nach der Quantenmechanik". In: *Zeitschrift für Physik* 44.6-7 (1927), pp. 455–472.

[84] Trygve Helgaker, Poul Jørgensen, and Jeppe Olsen. *Molecular Electronic Structure Theory*. West Sussex, England: John Wiley & Sons, Ltd., 2000.

[85] Jochen Heyd, Gustavo E. Scuseria, and Matthias Ernzerhof. "Erratum: hybrid functionals based on a screened Coulomb potential - J. Chem. Phys. 118, 8207 (2003)". In: *The Journal of Chemical Physics* 124.21 (June 2006), p. 219906. ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.2204597. URL: http://dx.doi.org/10.1063/1.2204597.

[86] Jochen Heyd, Gustavo E. Scuseria, and Matthias Ernzerhof. "Hybrid functionals based on a screened Coulomb potential". In: *The Journal of Chemical Physics* 118.18 (May 2003), pp. 8207–8215. ISSN: 0021-9606, 1089-7690. DOI: 10.1063/1.1564060. URL: http://dx.doi.org/10.1063/1.1564060.

[87] Jochen Heyd, Gustavo E. Scuseria, and Matthias Ernzerhof. "Hybrid functionals based on a screened Coulomb potential". In: *The Journal of Chemical Physics* 118.18 (2003), pp. 8207–8215. DOI: 10.1063/1.1564060. eprint: https://doi.org/10.1063/1.1564060. URL: https://doi.org/10.1063/1.1564060.

[88] Burak Himmetoglu et al. "Hubbard-corrected DFT energy functionals: The LDA+U description of correlated systems". In: *International Journal of Quantum Chemistry* 114.1 (2014), pp. 14–49. DOI: 10.1002/qua.24521. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/qua.24521. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/qua.24521.

[89] P. Hohenberg and W. Kohn. "Inhomogeneous Electron Gas". In: *Phys. Rev.* 136 (3B Nov. 1964), B864–B871. DOI: 10.1103/PhysRev.136.B864. URL: https://link.aps.org/doi/10.1103/PhysRev.136.B864.

[90] J.-W. van der Horst et al. "Ab initio prediction of the electronic and optical excitations in polythiophene: Isolated chains versus bulk polymer". In: *Phys. Rev. B* 61 (23 June 2000), pp. 15817–15826. DOI: 10.1103/PhysRevB.61.15817. URL: https://link.aps.org/doi/10.1103/PhysRevB.61.15817.

[91] A R Hoy and P R Bunker. "A precise solution of the rotation bending Schrödinger equation for a triatomic molecule with application to the water molecule". In: *J. Mol. Spectrosc.* 74.1 (1979), pp. 1–8.

[92] J. Hubbard and Brian Hilton Flowers. "Electron correlations in narrow energy bands". In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 276.1365 (1963), pp. 238–257. DOI: 10.1098/rspa.1963.0204. eprint: https://royalsocietypublishing.org/doi/pdf/10.1098/rspa.1963.0204. URL: https://royalsocietypublishing.org/doi/abs/10.1098/rspa.1963.0204.

[93] Bernard Huron, JP Malrieu, and P Rancurel. "Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wavefunctions". In: *The Journal of Chemical Physics* 58.12 (1973), pp. 5745–5759.

[94] Mark S. Hybertsen and Steven G. Louie. "Electron correlation in semiconductors and insulators: Band gaps and quasiparticle energies". In: *Phys. Rev. B* 34.8 (Oct. 1986), pp. 5390–5413. ISSN: 0163-1829. DOI: 10.1103/physrevb.34.5390. URL: https://doi.org/10.1103/physrevb.34.5390.

[95] Mark S. Hybertsen and Steven G. Louie. "First-Principles Theory of Quasiparticles: Calculation of Band Gaps in Semiconductors and Insulators". In: *Phys. Rev. Lett.* 55.13 (Sept. 1985), pp. 1418–1421. ISSN: 0031-9007. DOI: 10.1103/physrevlett.55.1418. URL: https://doi.org/10.1103/physrevlett.55.1418.

[96] R N Iskenderov et al. In: *Fiz. Tverd. Tela (Leningrad)* 10 (1968), p. 2573.

[97] Manish Jain, James R. Chelikowsky, and Steven G. Louie. "Reliability of Hybrid Functionals in Predicting Band Gaps". In: *Phys. Rev. Lett.* 107 (21 Nov. 2011), p. 216806. DOI: 10.1103/PhysRevLett.107.216806. URL: https://link.aps.org/doi/10.1103/PhysRevLett.107.216806.

[98] Manish Jain, James R Chelikowsky, and Steven G Louie. "Reliability of hybrid functionals in predicting band gaps". In: *Physical review letters* 107.21 (2011), p. 216806.

[99] Hans Jørgen Aa Jensen et al. "Accurate calculations of the dynamic dipole polarizability of N2. A multiconfigurational linear response study using restricted active space (RAS) wavefunctions". In: *Chemical physics letters* 162.4-5 (1989), pp. 355–360.

[100] Hong Jiang et al. "First-principles modeling of localized d states with the GW@LDA+U approach". In: *Phys. Rev. B* 82 (4 July 2010), p. 045108. DOI: 10.1103/PhysRevB.82.045108. URL: https://link.aps.org/doi/10.1103/PhysRevB.82.045108.

[101] Branko Jursic and Zoran Zdravkovski. "DFT study of the Diels–Alder reactions between ethylene with buta-1, 3-diene and cyclopentadiene". In: *Journal of the Chemical Society, Perkin Transactions 2* 6 (1995), pp. 1223–1226.

[102] Tosio Kato. "On the eigenfunctions of many-particle systems in quantum mechanics". In: *Communications on Pure and Applied Mathematics* 10.2 (1957), pp. 151–177.

[103] Alexander Khein. "Analysis of separable nonlocal pseudopotentials". In: *Phys. Rev. B* 51 (23 June 1995), pp. 16608–16614. DOI: 10.1103/PhysRevB.51.16608. URL: https://link.aps.org/doi/10.1103/PhysRevB.51.16608.

[104] Jeongnim Kim et al. "QMCPACK : An open source ab initio Quantum Monte Carlo package for the electronic structure of atoms, molecules, and solids". In: *J. Phys. Condens. Matter* 30 (2018), p. 195901. DOI: 10.1088/1361-648X/aab9c3.

[105] N. Kirova, S. Brazovskii, and A.R. Bishop. "A systematic theory for optical properties of phenylene-based polymers". In: *Synthetic Metals* 100.1 (1999), pp. 29–53. ISSN: 0379-6779. DOI: https://doi.org/10.1016/S0379-6779(98)00174-X. URL: http://www.sciencedirect.com/science/article/pii/S037967799800174X.

[106] Leonard Kleinman and D. M. Bylander. "Efficacious Form for Model Pseudopotentials". In: *Phys. Rev. Lett.* 48 (20 May 1982), pp. 1425–1428. DOI: 10.1103/PhysRevLett.48.1425. URL: https://link.aps.org/doi/10.1103/PhysRevLett.48.1425.

[107] Leonard Kleinman and DM Bylander. "Efficacious form for model pseudopotentials". In: *Physical Review Letters* 48.20 (1982), p. 1425.

[108] Walter Kohn and Lu Jeu Sham. "Self-consistent equations including exchange and correlation effects". In: *Physical review* 140.4A (1965), A1133.

[109] Stephen R Langhoff and Ernest R Davidson. "Configuration interaction calculations on the nitrogen molecule". In: *International Journal of Quantum Chemistry* 8.1 (1974), pp. 61–72.

[110] Stephan Lany. "Band-structure calculations for the 3 d transition metal oxides in G W". In: *Physical Review B* 87.8 (2013), p. 085112.

[111] Ji-Woo Lee, Lubos Mitas, and Lucas K. Wagner. *Quantum Monte Carlo study of MnO solid*. 2004. eprint: arXiv:cond-mat/0411247.

[112] Y. Lee and R. J. Needs. "Core-polarization potentials for Si and Ti". In: *Phys. Rev. B* 67 (3 Jan. 2003), p. 035121. DOI: 10.1103/PhysRevB.67.035121. URL: https://link.aps.org/doi/10.1103/PhysRevB.67.035121.

[113] G. Leising. "Anisotropy of the optical constants of pure and metallic polyacetylene". In: *Phys. Rev. B* 38 (15 Nov. 1988), pp. 10313–10322. DOI: 10.1103/PhysRevB.38.10313. URL: https://link.aps.org/doi/10.1103/PhysRevB.38.10313.

[114] Junhao Li et al. "Fast semistochastic heat-bath configuration interaction". In: *The Journal of chemical physics* 149.21 (2018), p. 214110.

[115] J Linnanto, JEI Korppi-Tommola, and VM Helenius. "Electronic States, Absorption Spectrum and Circular Dichroism Spectrum of the Photosynthetic Bacterial LH2 Antenna of Rhodopseudomonas a cidophila as Predicted by Exciton Theory and Semiempirical Calculations". In: *The Journal of Physical Chemistry B* 103.41 (1999), pp. 8739–8750.

[116] STEVEN G Louie. "First-principles theory of electron excitation energies in solids, surfaces, and defects". In: *Topics in Computational Materials Science* 96 (1998).

[117] Per Åke Malmqvist, Alistair Rendell, and Bjoern O Roos. "The restricted active space self-consistent-field method, implemented with a split graph unitary group approach". In: *Journal of Physical Chemistry* 94.14 (1990), pp. 5477–5482.

[118] Richard L. Martin. "Natural transition orbitals". In: *The Journal of Chemical Physics* 118.11 (2003), pp. 4775–4777. DOI: 10.1063/1.1558471. eprint: https://doi.org/10.1063/1.1558471. URL: https://doi.org/10.1063/1.1558471.

[119] Richard M. Martin. "Pseudopotentials". In: *Electronic Structure: Basic Theory and Practical Methods*. Cambridge University Press, 2004, pp. 204–232. DOI: 10.1017/CBO9780511805769.013.

[120] Walter E Massey and Chia-Wei Woo. "Variational Calculations on Liquid He 4 and He 3". In: *Physical Review* 164.1 (1967), p. 256.

[121] R.D. Mattuck. *A Guide to Feynman Diagrams in the Many-body Problem*. Dover Books on Physics Series. Dover Publications, Incorporated, 1976. ISBN: 9780486670478. URL: http://books.google.co.uk/books?id=pe-v8zfxE68C.

[122] Monica Mazzolini et al. "The phototransduction machinery in the rod outer segment has a strong efficacy gradient". In: *Proc. Natl. Acad. Sci. U.S.A.* 112.20 (2015), E2715–E2724. ISSN: 0027-8424. DOI: 10.1073/pnas.1423162112. eprint: http://www.pnas.org/content/112/20/E2715.full.pdf. URL: http://www.pnas.org/content/112/20/E2715.

[123] Yu Meng et al. "When density functional approximations meet iron oxides". In: *Journal of chemical theory and computation* 12.10 (2016), pp. 5132–5144.

[124] Albert Messiah and Leonard Isaac Schiff. *Quantum mechanics*. Vol. 643. McGraw-Hill College, 1968.

[125] Richard P Messmer. "On a variational method for determining excited state wave functions". In: *Theor. Chim. Acta.* 14.4 (1969), pp. 319–328.

[126] Nicholas Metropolis et al. "Equation of state calculations by fast computing machines". In: *The journal of chemical physics* 21.6 (1953), pp. 1087–1092.

[127] Miguel A. Morales et al. "Multideterminant Wave Functions in Quantum Monte Carlo". In: *J. Chem. Theory Comput.* 8.7 (2012), pp. 2181–2188. DOI: 10.1021/ct3003404.

[128] Paula Mori-Sánchez, Aron J Cohen, and Weitao Yang. "Localization and delocalization errors in density functional theory and implications for band-gap prediction". In: *Physical review letters* 100.14 (2008), p. 146401.

[129] B. Morosin. "Exchange Striction Effects in MnO and MnS". In: *Phys. Rev. B* 1 (1 Jan. 1970), pp. 236–243. DOI: 10.1103/PhysRevB.1.236. URL: https://link.aps.org/doi/10.1103/PhysRevB.1.236.

[130] Jules W Moskowitz et al. "A new look at correlation energy in atomic and molecular systems. II. The application of the Green's function Monte Carlo method to LiH". In: *The Journal of chemical physics* 77.1 (1982), pp. 349–355.

[131] JW Moskowitz and KE Schmidt. "Correlated Monte Carlo wave functions for some cations and anions of the first row atoms". In: *The Journal of chemical physics* 97.5 (1992), pp. 3382–3385.

[132] N F Mott. "The Basis of the Electron Theory of Metals, with Special Reference to the Transition Metals". In: *Proceedings of the Physical Society. Section A* 62.7 (July 1949), pp. 416–422. DOI: 10.1088/0370-1298/62/7/303. URL: https://doi.org/10.1088%5C%2F0370-1298%5C%2F62%5C%2F7%5C%2F303.

[133] Wolfgang Müller, Joachim Flesch, and Wilfried Meyer. "Treatment of intershell correlation effects in ab initio calculations by use of core polarization potentials. Method and application to alkali and alkaline earth atoms". In: *The Journal of Chemical Physics* 80.7 (1984), pp. 3297–3310. DOI: 10.1063/1.447083. eprint: https://doi.org/10.1063/1.447083. URL: https://doi.org/10.1063/1.447083.

[134] J Muscat, A Wander, and NM Harrison. "On the prediction of band gaps from hybrid functional theory". In: *Chemical Physics Letters* 342.3-4 (2001), pp. 397–401.

[135] Saeed Nasiri and Mansour Zahedi. "Quantum Monte Carlo study of ground and first excited state of C, N, O, F, and Ne atoms using Slater-Jastrow-Backflow wave function". In: *International Journal of Quantum Chemistry* 120.11 (2020), e26187.

[136] Almeria Natansohn and Paul Rochon. "Photoinduced Motions in Azo-Containing Polymers". In: *Chem. Rev.* 102.11 (2002), pp. 4139–4176. DOI: 10.1021/cr970155y.

[137] RK Nesbet. "Configuration interaction in orbital theories". In: *Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences* 230.1182 (1955), pp. 312–321.

[138] Eric Neuscamman. "The Jastrow antisymmetric geminal power in Hilbert space: Theory, benchmarking, and application to a novel transition state". In: *The Journal of chemical physics* 139.19 (2013), p. 194105.

[139] YM Niquet and Xavier Gonze. "Band-gap energy in the random-phase approximation to density-functional theory". In: *Physical Review B* 70.24 (2004), p. 245115.

[140] Jeppe Olsen. "The CASSCF method: A perspective and commentary". In: *International Journal of Quantum Chemistry* 111.13 (2011), pp. 3267–3272.

[141] Jeppe Olsen et al. "Determinant based configuration interaction algorithms for complete and restricted configuration interaction spaces". In: *The Journal of chemical physics* 89.4 (1988), pp. 2185–2192.

[142] Giovanni Onida, Lucia Reining, and Angel Rubio. "Electronic excitations: density-functional versus many-body Green's-function approaches". In: *Rev. Mod. Phys.* 74 (2 June 2002), pp. 601–659. DOI: 10.1103/RevModPhys.74.601. URL: https://link.aps.org/doi/10.1103/RevModPhys.74.601.

[143] Sokrates T. Pantelides, Daniel J. Mickish, and A. Barry Kunz. "Electronic structure and properties of magnesium oxide". In: *Phys. Rev. B* 10 (12 Dec. 1974), pp. 5203–5212. DOI: 10.1103/PhysRevB.10.5203. URL: https://link.aps.org/doi/10.1103/PhysRevB.10.5203.

[144] J Park et al. "Optical Properties of Transition-metal Oxides of MnO and Fe0. 925O Crystals Studied with Spectroscopic Ellipsometry and Raman Spectroscopy". In: *New Physics* 63 (2013), pp. 818–825.

[145] Robert G Parr. "Density functional theory of atoms and molecules". In: *Horizons of quantum chemistry*. Springer, 1980, pp. 5–15.

[146] Wolfgang Pauli. "Über das Wasserstoffspektrum vom Standpunkt der neuen Quantenmechanik". In: *Zeitschrift für Physik* 36.5 (1926), pp. 336–363.

[147] Wolfgang Pauli. "Über den Zusammenhang des Abschlusses der Elektronengruppen im Atom mit der Komplexstruktur der Spektren". In: *Zeitschrift für Physik* 31.1 (1925), pp. 765–783.

[148] Linus Pauling. "The nature of the chemical bond. III. The transition from one extreme bond type to another". In: *Journal of the American Chemical Society* 54.3 (1932), pp. 988–1003.

[149] R. E. Peierls. "The Commutation Laws of Relativistic Field Theory". In: *Proceedings of the Royal Society of London Series A* 214.1117 (Aug. 1952), pp. 143–157. DOI: 10.1098/rspa.1952.0158.

[150] John P. Perdew, Matthias Ernzerhof, and Kieron Burke. "Rationale for mixing exact exchange with density functional approximations". In: *The Journal of Chemical Physics* 105.22 (1996), pp. 9982–9985. DOI: 10.1063/1.472933. eprint: https://doi.org/10.1063/1.472933. URL: https://doi.org/10.1063/1.472933.

[151] John P. Perdew and Mel Levy. "Physical Content of the Exact Kohn-Sham Orbital Energies: Band Gaps and Derivative Discontinuities". In: *Phys. Rev. Lett.* 51 (20 Nov. 1983), pp. 1884–1887. DOI: 10.1103/PhysRevLett.51.1884. URL: https://link.aps.org/doi/10.1103/PhysRevLett.51.1884.

[152] John P. Perdew and Yue Wang. "Accurate and simple analytic representation of the electron-gas correlation energy". In: *Phys. Rev. B* 45.23 (June 1992), pp. 13244–13249. ISSN: 0163-1829, 1095-3795. DOI: 10.1103/physrevb.45.13244. URL: http://dx.doi.org/10.1103/physrevb.45.13244.

[153] John P. Perdew et al. "Density-Functional Theory for Fractional Particle Number: Derivative Discontinuities of the Energy". In: *Phys. Rev. Lett.* 49 (23 Dec. 1982), pp. 1691–1694. DOI: 10.1103/PhysRevLett.49.1691. URL: https://link.aps.org/doi/10.1103/PhysRevLett.49.1691.

[154] John P Perdew and Mel Levy. "Physical content of the exact Kohn-Sham orbital energies: band gaps and derivative discontinuities". In: *Physical Review Letters* 51.20 (1983), p. 1884.

[155] Sergio D. Pineda Flores and Eric Neuscamman. "Excited State Specific Multi-Slater Jastrow Wave Functions". In: *The Journal of Physical Chemistry A* 123.8 (2019). PMID: 30702890, pp. 1487–1497. DOI: 10.1021/acs.jpca.8b10671. eprint: https://doi.org/10.1021/acs.jpca.8b10671. URL: https://doi.org/10.1021/acs.jpca.8b10671.

[156] Sergio D Pineda Flores et al. "Bio-Inspired electroactive organic molecules for aqueous redox flow batteries. 1. Thiophenoquinones". In: *The Journal of Physical Chemistry C* 119.38 (2015), pp. 21800–21809.

[157] John A Pople, Martin Head-Gordon, and Krishnan Raghavachari. "Quadratic configuration interaction. A general technique for determining electron correlation energies". In: *The Journal of chemical physics* 87.10 (1987), pp. 5968–5975.

[158] Duminda S Ranasinghe et al. "Does the ionization potential condition employed in QTP functionals mitigate the self-interaction error?" In: *The Journal of chemical physics* 146.3 (2017), p. 034102.

[159] Peter J Reynolds et al. "Fixed-node quantum Monte Carlo for moleculesa) b)". In: *The Journal of Chemical Physics* 77.11 (1982), pp. 5593–5603.

[160] Paul J. Robinson, Sergio D. Pineda Flores, and Eric Neuscamman. "Excitation variance matching with limited configuration interaction expansions in variational Monte Carlo". In: *The Journal of Chemical Physics* 147.16 (2017), p. 164114. DOI: 10.1063/1.5008743. eprint: https://doi.org/10.1063/1.5008743. URL: https://doi.org/10.1063/1.5008743.

[161] Paul J Robinson, Sergio D Pineda Flores, and Eric Neuscamman. "Excitation variance matching with limited configuration interaction expansions in variational Monte Carlo". In: *J. Chem. Phys.* 147.16 (2017), p. 164114.

[162] C. Rödl et al. "Ab initio theory of excitons and optical properties for spin-polarized systems: Application to antiferromagnetic MnO". In: *Phys. Rev. B* 77 (18 May 2008), p. 184408. DOI: 10.1103/PhysRevB.77.184408. URL: https://link.aps.org/doi/10.1103/PhysRevB.77.184408.

[163] Michael Rohlfing and Steven G. Louie. "Optical Excitations in Conjugated Polymers". In: *Phys. Rev. Lett.* 82 (9 Mar. 1999), pp. 1959–1962. DOI: 10.1103/PhysRevLett.82.1959. URL: https://link.aps.org/doi/10.1103/PhysRevLett.82.1959.

[164] Björn O. Roos and Kerstin Andersson. "Multiconfigurational perturbation theory with level shift — the Cr2 potential revisited". In: *Chem. Phys. Lett.* 245.2 (1995), pp. 215–223. ISSN: 0009-2614. DOI: https://doi.org/10.1016/0009-2614(95)01010-7. URL: http://www.sciencedirect.com/science/article/pii/0009261495010107.

[165] Björn O Roos and Kerstin Andersson. "Multiconfigurational perturbation theory with level shift—the Cr2 potential revisited". In: *Chemical physics letters* 245.2-3 (1995), pp. 215–223.

[166] Björn O Roos, Peter R Taylor, and Per EM Sigbahn. "A complete active space SCF method (CASSCF) using a density matrix formulated super-CI approach". In: *Chemical Physics* 48.2 (1980), pp. 157–173.

[167] Bjorn O Roos et al. "Multiconfigurational perturbation theory: Applications in electronic spectroscopy". In: *Advances in chemical physics, vol xciii* 93 (1996), pp. 219–331.

[168] Clemens Carel Johannes Roothaan. "New developments in molecular orbital theory". In: *Reviews of modern physics* 23.2 (1951), p. 69.

[169] Robert Roth. "Importance truncation for large-scale configuration interaction approaches". In: *Physical Review C* 79.6 (2009), p. 064324.

[170] W. L. Roth. "Magnetic Structures of MnO, FeO, CoO, and NiO". In: *Phys. Rev.* 110 (6 June 1958), pp. 1333–1341. DOI: 10.1103/PhysRev.110.1333. URL: https://link.aps.org/doi/10.1103/PhysRev.110.1333.

[171] Mario Russo et al. "Polymerization and preliminary X-ray analysis of 1,3,5,7-tetrathiocane (tetra-thioformaldehyde)". In: *J. Polym. Sci. B* 3.6 (1965), pp. 501–504. DOI: 10.1002/pol.1965.110030616. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pol.1965.110030616. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/pol.1965.110030616.

[172] Espen Sagvolden and John P. Perdew. "Discontinuity of the exchange-correlation potential: Support for assumptions used to find it". In: *Phys. Rev. A* 77 (1 Jan. 2008), p. 012517. DOI: 10.1103/PhysRevA.77.012517. URL: https://link.aps.org/doi/10.1103/PhysRevA.77.012517.

[173] U Salzner et al. "Accurate method for obtaining band gaps in conducting polymers using a DFT/hybrid approach". In: *The Journal of Physical Chemistry A* 102.15 (1998), pp. 2572–2578.

[174] Georgy Samsonidze, Cheol-Hwan Park, and Boris Kozinsky. "Insights and challenges of applying theGWmethod to transition metal oxides". In: *Journal of Physics: Condensed Matter* 26.47 (Oct. 2014), p. 475501. DOI: 10.1088/0953-8984/26/47/475501. URL: https://doi.org/10.1088%2F0953-8984%2F26%2F47%2F475501.

[175] Juan A. Santana et al. "Structural stability and defect energetics of ZnO from diffusion quantum Monte Carlo". In: *The Journal of Chemical Physics* 142.16 (2015), p. 164705. DOI: 10.1063/1.4919242. eprint: https://doi.org/10.1063/1.4919242. URL: https://doi.org/10.1063/1.4919242.

[176] Brett M Savoie et al. "Reassessing the use of one-electron energetics in the design and characterization of organic photovoltaics". In: *Physical Chemistry Chemical Physics* 15.13 (2013), pp. 4538–4547.

[177] Friedemann Schautz, Francesco Buda, and Claudia Filippi. "Excitations in photoactive molecules from quantum Monte Carlo". In: *J. Chem. Phys.* 121.12 (2004), pp. 5836–5844. DOI: 10.1063/1.1777212.

[178] Friedemann Schautz and Stephen Fahy. "Optimization of configuration interaction coefficients in multideterminant Jastrow–Slater wave functions". In: *The Journal of chemical physics* 116.9 (2002), pp. 3533–3539.

[179] Friedemann Schautz and Claudia Filippi. "Optimized Jastrow–Slater wave functions for ground and excited states: application to the lowest states of ethene". In: *The Journal of chemical physics* 120.23 (2004), pp. 10931–10941.

[180] Joshua A Schiller, Lucas K Wagner, and Elif Ertekin. "Phase stability and properties of manganese oxide polymorphs: Assessment and insights from diffusion Monte Carlo". In: *Phys. Rev. B* 92.23 (2015), p. 235209.

[181] U Schönberger and F Aryasetiawan. "Bulk and surface electronic structures of MgO". In: *Physical Review B* 52.12 (1995), p. 8788.

[182] JB Schriber and FA Evangelista. "Communication: An adaptive configuration interaction approach for strongly correlated electrons with tunable accuracy." In: *J. Chem. Phys.* 144.16 (2016), p. 161106.

[183] Erwin Schrödinger. "An undulatory theory of the mechanics of atoms and molecules". In: *Physical review* 28.6 (1926), p. 1049.

[184] A. Schrön, C. Rödl, and F. Bechstedt. "Energetic stability and magnetic properties of MnO in the rocksalt, wurtzite, and zinc-blende structures: Influence of exchange and correlation". In: *Phys. Rev. B* 82 (16 Oct. 2010), p. 165109. DOI: 10.1103/PhysRevB.82.165109. URL: https://link.aps.org/doi/10.1103/PhysRevB.82.165109.

[185] M. J. van Setten, F. Weigend, and F. Evers. "The GW-Method for Quantum Chemistry Applications: Theory and Implementation". In: *Journal of Chemical Theory and Computation* 9.1 (2013). PMID: 26589026, pp. 232–246. DOI: 10.1021/ct300648t. eprint: https://doi.org/10.1021/ct300648t. URL: https://doi.org/10.1021/ct300648t.

[186] Sason Shaik et al. "P450 Enzymes: Their Structure, Reactivity, and Selectivity Modeled by QM/MM Calculations". In: *Chemical reviews* 110.2 (2010), pp. 949–1017.

[187] L. J. Sham and M. Schlüter. "Density-Functional Theory of the Energy Gap". In: *Phys. Rev. Lett.* 51 (20 Nov. 1983), pp. 1888–1891. DOI: 10.1103/PhysRevLett.51.1888. URL: https://link.aps.org/doi/10.1103/PhysRevLett.51.1888.

[188] Yihan Shao et al. "Advances in molecular quantum chemistry contained in the Q-Chem 4 program package". In: *Molecular Physics* 113.2 (2015), pp. 184–215.

[189] Sandeep Sharma et al. "Semistochastic heat-bath configuration interaction method: Selected configuration interaction with semistochastic perturbation theory". In: *Journal of chemical theory and computation* 13.4 (2017), pp. 1595–1604.

[190] Jacqueline A. R. Shea and Eric Neuscamman. "Size Consistent Excited States via Algorithmic Transformations between Variational Principles". In: *J. Chem. Theory Comput.* 13 (2017), p. 6078.

[191] Jacqueline AR Shea and Eric Neuscamman. "Communication: A mean field platform for excited state quantum chemistry". In: *J. Chem. Phys.* 149.8 (2018), p. 081101.

[192] K. Shimamura et al. "Crystal structure of trans-polyacetylene". In: *Die Makromolekulare Chemie, Rapid Communications* 2.8 (1981), pp. 473–480. DOI: 10.1002/marc.1981.030020801. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/marc.1981.030020801. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/marc.1981.030020801.

[193] Hideki Shirakawa et al. "Synthesis of electrically conducting organic polymers: halogen derivatives of polyacetylene, (CH)". In: *J. Chem. Soc., Chem. Commun.* (16 1977), pp. 578–580. DOI: 10.1039/C39770000578. URL: http://dx.doi.org/10.1039/C39770000578.

[194] Eric L. Shirley and Richard M. Martin. "GW quasiparticle calculations in atoms". In: *Phys. Rev. B* 47 (23 June 1993), pp. 15404–15412. DOI: 10.1103/PhysRevB.47.15404. URL: https://link.aps.org/doi/10.1103/PhysRevB.47.15404.

[195] Eric L. Shirley and Richard M. Martin. "Many-body core-valence partitioning". In: *Phys. Rev. B* 47 (23 June 1993), pp. 15413–15427. DOI: 10.1103/PhysRevB.47.15413. URL: https://link.aps.org/doi/10.1103/PhysRevB.47.15413.

[196] Eric L. Shirley, Lubo š Mitá š, and Richard M. Martin. "Core-valence partitioning and quasiparticle pseudopotentials". In: *Phys. Rev. B* 44 (7 Aug. 1991), pp. 3395–3398. DOI: 10.1103/PhysRevB.44.3395. URL: https://link.aps.org/doi/10.1103/PhysRevB.44.3395.

[197] Cleiton M. da Silva et al. "Schiff bases: A short review of their antimicrobial activities". In: *J. Adv. Res.* 2.1 (2011), pp. 1–8. ISSN: 2090-1232. DOI: https://doi.org/10.1016/j.jare.2010.05.004. URL: http://www.sciencedirect.com/science/article/pii/S2090123210000603.

[198] James E T Smith et al. "Cheap and near exact CASSCF with large active spaces". In: *J. Chem. Theory Comput.* 13.11 (2017), pp. 5468–5478.

[199] V. S. Stepanyuk et al. "Electronic Structure and Optical Properties of Mgo Band Structure Calculation and Cluster Model". In: *physica status solidi (b)* 155.1 (1989), pp. 179–184. DOI: 10.1002/pssb.2221550116. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1002/pssb.2221550116. URL: https://onlinelibrary.wiley.com/doi/abs/10.1002/pssb.2221550116.

[200] Attila Szabo and Neil S. Ostlund. *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*. Mineola, N.Y.: Dover Publications, 1996.

[201] Teresa Tamayo-Mendoza et al. "Automatic Differentiation in Quantum Chemistry with Applications to Fully Variational Hartree–Fock". In: *ACS Cent. Sci.* 4.5 (2018), p. 559.

[202] Ivano Tavernelli, Ute F. Röhrig, and Ursula Rothlisberger. "Molecular dynamics in electronically excited states using time-dependent density functional theory". In: *Mol. Phys.* 103.6-8 (2005), pp. 963–981. DOI: 10.1080/00268970512331339378.

[203] *The Nobel Prize in Chemistry 2000*. URL: https://www.nobelprize.org/prizes/chemistry/2000/summary/.

[204] Murilo L. Tiago, Michael Rohlfing, and Steven G. Louie. "Bound excitons and optical properties of bulk trans-polyacetylene". In: *Phys. Rev. B* 70 (19 Nov. 2004), p. 193204. DOI: 10.1103/PhysRevB.70.193204. URL: https://link.aps.org/doi/10.1103/PhysRevB.70.193204.

[205] William W. Tipton, Neil D. Drummond, and Richard G. Hennig. "Importance of high-angular-momentum channels in pseudopotentials for quantum Monte Carlo". In: *Phys. Rev. B* 90 (12 Sept. 2014), p. 125110. DOI: 10.1103/PhysRevB.90.125110. URL: https://link.aps.org/doi/10.1103/PhysRevB.90.125110.

[206] Julien Toulouse, Roland Assaraf, and Cyrus J. Umrigar. "Chapter Fifteen - Introduction to the Variational and Diffusion Monte Carlo Methods". In: *Electron Correlation in Molecules – ab initio Beyond Gaussian Quantum Chemistry*. Ed. by Philip E. Hoggan and Telhat Ozdogan. Vol. 73. Advances in Quantum Chemistry. Academic Press, 2016, pp. 285–314. DOI: https://doi.org/10.1016/bs.aiq.2015.07.003. URL: http://www.sciencedirect.com/science/article/pii/S0065327615000386.

[207] Julien Toulouse and C. J. Umrigar. "Full optimization of Jastrow–Slater wave functions with application to the first-row atoms and homonuclear diatomic molecules". In: *J. Chem. Phys.* 128.17 (2008), p. 174101. DOI: 10.1063/1.2908237.

[208] Julien Toulouse and CJ Umrigar. "Full optimization of Jastrow–Slater wave functions with application to the first-row atoms and homonuclear diatomic molecules". In: *The Journal of chemical physics* 128.17 (2008), p. 174101.

[209] M. D. Towler et al. "Ab initio study of MnO and NiO". In: *Phys. Rev. B* 50 (8 Aug. 1994), pp. 5041–5054. DOI: 10.1103/PhysRevB.50.5041. URL: https://link.aps.org/doi/10.1103/PhysRevB.50.5041.

[210] J. P. Townsend et al. In: *submitted to PRB(R)* (2020).

[211] J R Trail. "Alternative sampling for variational quantum Monte Carlo." In: *Phys. Rev. E* 77 (Jan. 2008), p. 016704.

[212] J R Trail. "Heavy-tailed random error in quantum Monte Carlo." In: *Phys. Rev. E* 77 (Jan. 2008), p. 016703.

[213] Norman Troullier and Jose Luis Martins. "Efficient pseudopotentials for plane-wave calculations". In: *Physical review B* 43.3 (1991), p. 1993.

[214] Norm M Tubman et al. "A deterministic alternative to the full configuration interaction quantum Monte Carlo method". In: *The Journal of chemical physics* 145.4 (2016), p. 044112.

[215] CJ Umrigar et al. "Alleviation of the fermion-sign problem by optimization of many-body wave functions". In: *Phys. Rev. Lett.* 98.11 (2007), p. 110201.

[216] Ting Wang, Xiaojun Zhou, and Fan Wang. "Performance of the diffusion quantum Monte Carlo method with a single-Slater-Jastrow trial wavefunction using natural orbitals and density functional theory orbitals on atomization energies of the Gaussian-2 set". In: *The Journal of Physical Chemistry A* 123.17 (2019), pp. 3809–3817.

[217] H.-J. Werner et al. "Molpro: a general-purpose quantum chemistry program package". In: *WIREs Comput. Mol. Sci.* 2 (2012), p. 242. DOI: 10.1002/wcms.82.

[218] R. C. Whited, Christopher J. Flaten, and W. C. Walker. "Exciton thermoreflectance of MgO and CaO". In: *Solid State Communications* 13.11 (Dec. 1973), pp. 1903–1905. DOI: 10.1016/0038-1098(73)90754-0.

[219] Wikipedia, the free encyclopedia. *Pseudopotential*. [Online; accessed April 20, 2019]. 2019. URL: https://upload.wikimedia.org/wikipedia/commons/f/fb/Sketch_Pseudopotentials.png.

[220] M Wolf et al. "Solar cell efficiency and carrier multiplication in Si 1- x Ge x alloys". In: *Journal of applied physics* 83.8 (1998), pp. 4213–4221.

[221] Hai Xiao, Jamil Tahir-Kheli, and William A Goddard III. "Accurate band gaps for semiconductors from density functional theory". In: *The Journal of Physical Chemistry Letters* 2.3 (2011), pp. 212–217.

[222] Hong-Zhou Ye et al. "σ-SCF: A direct energy-targeting method to mean-field excited states". In: *J. Chem. Phys.* 147.21 (2017), p. 214104.

[223] J. Zaanen, G. A. Sawatzky, and J. W. Allen. "Band gaps and electronic structure of transition-metal compounds". In: *Phys. Rev. Lett.* 55 (4 July 1985), pp. 418–421. DOI: 10.1103/PhysRevLett.55.418. URL: https://link.aps.org/doi/10.1103/PhysRevLett.55.418.

[224] J. Zaanen and G.A. Sawatzky. "Systematics in band gaps and optical spectra of 3D transition metal compounds". In: *Journal of Solid State Chemistry* 88.1 (1990), pp. 8–27. ISSN: 0022-4596. DOI: https://doi.org/10.1016/0022-4596(90)90202-9. URL: http://www.sciencedirect.com/science/article/pii/0022459690902029.

[225] J. Zemann. "Crystal structures, 2nd edition. Vol. 1 by R. W. G. Wyckoff". In: *Acta Crystallographica* 18.1 (1965), pp. 139–139. DOI: 10.1107/S0365110X65000361. eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1107/S0365110X65000361. URL: https://onlinelibrary.wiley.com/doi/abs/10.1107/S0365110X65000361.

[226] Luning Zhao and Eric Neuscamman. "A Blocked Linear Method for Optimizing Large Parameter Sets in Variational Monte Carlo". In: *J. Chem. Theory Comput.* 13.6 (2017), pp. 2604–2611.

[227] Luning Zhao and Eric Neuscamman. "An efficient variational principle for the direct optimization of excited states". In: *J. Chem. Theory Comput.* 12 (2016), p. 3436.

[228] Luning Zhao and Eric Neuscamman. "Variational Excitations in Real Solids: Optical Gaps and Insights into Many-Body Perturbation Theory". In: *Phys. Rev. Lett.* 123 (3 July 2019), p. 036402. DOI: 10.1103/PhysRevLett.123.036402. URL: https://link.aps.org/doi/10.1103/PhysRevLett.123.036402.

[229] J Patrick Zobel, Juan J Nogueira, and Leticia González. "The IPEA dilemma in CASPT2". In: *Chemical science* 8.2 (2017), pp. 1482–1499.

Appendix A

Appendix for Aperiodic Systems

A.1 Computational Details

For quantum Monte Carlo calculations, all wave function parameters were optimized by variationally minimizing Ω using the ω-update scheme presented in Ref. [190]. Both the two-body and all one-body Jastrow factors took a Bspline form with a cutoff at 10 bohr and used 10 spline points. [104]

H_4

Our skew arrangement of four H atoms was chosen to remove all symmetry and to create a simple, small system in which strong correlation was present so that orbital relaxations in a small MSJ expansion would be expected to make a difference. We employed BFD effective core potentials and the corresponding VTZ basis [29] and placed the atoms at the positions given below in Angstroms. The configurations for the MSJ were chosen by using the 10 most important configurations from a ground state CASSCF (4e,10o) calculation.

H	0.0000000000	0.0000000000	0.0000000000
H	1.8897259877	0.0000000000	0.0000000000
H	0.0000000000	0.0000000000	2.8345889816
H	0.0000000000	0.0000000000	5.6691779632

H_2O

For our calculations on water we employed BFD effective core potentials with the VDZ basis [29] at the experimental equilbirum geometry [91] given in Angstroms below. The 100 configurations for the MSJ excited state singlet were chosen as the largest-weight configurations in the excited state of a two-state full-valence CIPSI calculation in the RHF orbital basis.

O	0.0000	0.0000	0.1173
H	0.0000	0.7572	-0.4692
H	0.0000	-0.7572	-0.4692

$\mathbf{CH_2S}$

For thioformaldehyde we used BFD effective core potentials with their VTZ basis [29] and the following geometry, in Angstroms.

S	-4.9615006425	2.6553412397	0.0000217073
C	-4.9017991394	1.0716634201	-0.0001062888
H	-5.5890742022	0.4742274771	0.5685871400
H	-4.1719760160	0.5275278631	-0.5685025585

$\left[\mathbf{C_3N_2O_2H_4Cl}\right]^-$

For our chlorine-to-π^* charge transfer system, we used the geometry given in Angstroms below that was arrived at via an MP2/cc-pVDZ geometry optimization in Molpro for the closed-shell anionic ground state. For all excitation energy evaluations, we employed BFD effective core potentials with the corresponding VDZ basis. [29]

N	-2.9058516510	0.0000000000	-1.4601212300
C	-1.6406693060	0.0000000000	-1.1581565780
C	-1.2646457880	0.0000000000	0.2633575220
C	0.0390698630	0.0000000000	0.6172500750
N	0.4054577830	0.0000000000	2.0353753730
O	1.6182721380	0.0000000000	2.2841014810
O	-0.4846283820	0.0000000000	2.8946943050
Cl	1.5876389750	0.0000000000	-2.3723480080
H	0.8777584070	0.0000000000	-0.1013594360
H	-2.0565965790	0.0000000000	1.0201589650
H	-0.7953016620	0.0000000000	-1.8742706390
H	-2.9621125330	0.0000000000	-2.4917011270

$\mathbf{CH_2NH}$

For formaldimine, we used BFD effective core potentials and their VTZ basis. [29] All active space methods were based on equally-weighted two-state state averaged CASSCF wave functions. A (2e,2o) active space ground state CASSCF geometry optimization was

performed with the constraint that the dihedral angle of the molecule remain fixed at 0, 45, or 90 degree.

	Torsion Coordinate 0°		
C	0.0000000000	0.0222705613	-0.6225060885
N	0.0000000000	-0.0820734375	0.6148349963
H	0.0000000000	0.9991273614	-1.0742328325
H	0.0000000000	-0.8651411291	-1.2314451255
H	0.0000000000	0.7411514791	1.1797292763
	Torsion Coordinate 45°		
C	-0.0079599003	0.0166001095	-0.6325884466
N	0.0416062329	-0.0648316530	0.6270535275
H	0.1330377950	0.9587507613	-1.1388276756
H	-0.1929029663	-0.8533530806	-1.2431487361
H	-0.4234569109	0.5977132175	1.2063798375
	Torsion Coordinate 90°		
C	-0.0144604887	0.0000000000	-0.6239939985
N	0.0700010603	0.0000000000	0.6267505389
H	-0.0579593274	0.8804123024	-1.2681443330
H	-0.0579593274	-0.8804123024	-1.2681443330
H	-0.6845247458	0.0000000000	1.2624877883

A.2 State Averaging Without Symmetry

While the state-averaging sensitivity we observed for $\left[C_3N_2O_2H_4Cl\right]^-$ in the main text was already worrisome, the situation would be drastically worse were this molecule not C_s symmetric. In that case, it is not clear that it is possible to accurately model the out-of-plane-Cl-3P $\rightarrow \pi^*$ excitation energy via standard state-averaged multi-reference methods. Indeed, when we did not exploit the symmetry and instead treated the molecule as if it were C_1, we were unable to find either a 2-state or 3-state state averaged CASSCF that contained the out-of-plane-Cl-3p $\rightarrow \pi^*$ excitation, despite trying numerous initial guesses and optimization methods. Either one or both of the two in-plane-Cl-3p $\rightarrow \pi^*$ excitations ended up being lower in energy, or, for some initial guesses, the orbitals optimized to be so favorable for states with ground-state-like charge distributions that none of the excitations turned out to involve moving charge away from the chlorine atom. It was only when we resorted to a 4-state state average CASSCF that we were able to find the desired state, which turned out to come out alongside the ground state and the two in-plane-Cl-3p $\rightarrow \pi^*$ excitations. In this case, three of the four states in the state average have a dipole greatly

different from that of the ground state, and so as one would expect the orbital optimization favored them at the expense of the ground state, resulting in a much too small CASSCF excitation energy, as seen in Table II of the main text. As real chemical environments such as protein superstructures or solvents typically remove symmetry, one realizes that these types of state averaging difficulties will be quite common when attempting to model charge transfer in large molecules and realistic environments.

Appendix B

Appendix for Periodic Systems

B.1 Computational Details

Below are the Mathematica notebooks containing the computational details for periodic system calculations.

Supplementary Information

Sergio Pineda Flores

All Jastrow parameters had a cut off radius equal to the Wigner-Seitz radius
one/two body jastrow had 8 parameters (spline points)
three body had 26 parameters (spline spoints)

MgO (HSE06 starting Orbitals, Ne core BFD ECP, kinetic energy cutoff = 320 Ry) Variance matched gaps and finite size corrections

supercell	grnd_enrgy	std_err	variance	std_err	num_dets	exctd_enrgy	std_err	variance	std_err	num_dets
1x1x1	−16.8231	0.000860339	0.325973	0.00275832	1	−16.7746	0.000845235	0.336216	0.00402565	73
2x1x1	−33.8813	0.0014402	0.759865	0.00593372	1	−33.7201	0.00221329	0.761749	0.00744455	3
2x2x1	−68.0515	0.00494328	1.62435	0.0165583	1	−67.818	0.00203203	1.63002	0.00768934	1153

2 | test.nb

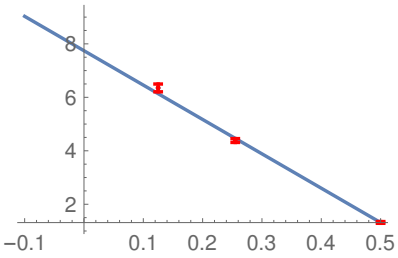
```
In[ ]:= data = {
          {0.5, 1.3204296800356},
          {0.255, 4.38405636282429},
          {0.125, 6.35192910364676}
         };
        errors = {0.032818862339022,
          0.07185461381465,
          0.145435011293214};
        withError = Transpose[{data[[All, 1]], data[[All, 2]], errors}];
        nlm = NonlinearModelFit[data, m x + b, {m, b}, {x}, Weights → 1 / errors^2]
        nlm[{"BestFit", "ParameterTable"}]
        Needs["ErrorBarPlots`"]
        Show[Plot[nlm[x], {x, -0.1, 0.5}], ListPlot[data],
         ErrorListPlot[withError, PlotStyle → Red]]
```

```
Out[ ]= FittedModel[ 7.74333 - 12.8575 x ]
```

```
Out[ ]= {7.74333 - 12.8575 x,
```

	Estimate	Standard Error	t-Statistic	P-Value
m	-12.8575	0.488358	-26.328	0.0241687
b	7.74333	0.223734	34.6096	0.0183892

}



PAE (HF starting Orbitals, BFD ECP, kinetic energy cutoff = 200 Ry) Variance matched gaps and finite size corrections

supercell	grnd_enrgy	std_err	variance	std_err	num_dets	exctd_enrgy	std_err	variance	std_err	num_dets
1x1x1	-24.6244	0.00155823	0.834081	0.00310832	1	-24.4035	0.00154927	0.894865	0.00419901	4
1x1x2	-49.8718	0.00306322	0.941018	0.00885931	1	-49.7364	0.00309091	0.949408	0.00750747	4
1x1x4	-99.9125	0.00431821	1.87347	0.0126489	1	-99.8185	0.00434812	1.87917	0.0139447	4

In[ ]:=
```
data = {
    {0.125, 6.01041372130406},
    {0.0625, 3.68398850223182},
    {0.03125, 2.55638266566293}
   };
errors = {0.059792948751601,
    0.11841508575809,
    0.166752843045092};
withError = Transpose[{data[[All, 1]], data[[All, 2]], errors}];
nlm = NonlinearModelFit[data, m x + b, {m, b}, {x}, Weights → 1 / errors^2]
nlm[{"BestFit", "ParameterTable"}]
Needs["ErrorBarPlots`"]
Show[Plot[nlm[x], {x, -0.01, 0.14}], ListPlot[data],
 ErrorListPlot[withError, PlotStyle → Red]]
```

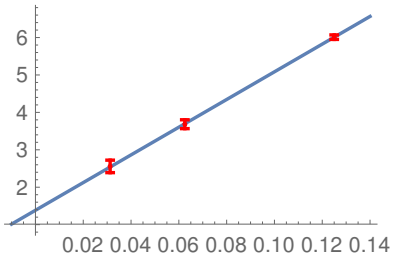
Out[ ]=
```
FittedModel[ 1.38348 + 37.007 x ]
```

Out[ ]=

1.38348 + 37.007 x

	Estimate	Standard Error	t-Statistic	P-Value
m	37.007	0.21962	168.505	0.003778
b	1.38348	0.0241623	57.2579	0.0111173

Out[ ]=



Gamma -> Gamma transition of MnO

VMC MnO HSE06 starting orbitals (RRKJ ECP for Mn and Shin ECP for O, kinetic energy cutoff = 400 Ry)

supercell	grnd_enrgy	std_err	variance	std_err	num_dets	exctd_enrgy	std_err	variance	std_err	num_dets
2x2x2 (chemical unit cell)	−960.557	0.00138	19.4411	0.0251	79	−960.351	0.00099	19.4306	0.04982	14

In[ ]:=
```
(* optical gap in eV *)
(-960.351282 + 960.556514) * 27.2114
```

Out[ ]=
```
5.58465
```

In[ ]:=
```
(* standard error of gap in eV*)
(√(0.00138^2 + 0.00099^2)) * 27.2114
```

Out[ ]=
```
0.0462153
```

16 atom unit cell

ATOMIC_POSITIONS crystal
Mn1 0.00 0.00 0.00
Mn2 0.50 0.50 0.50
Mn2 0.50 0.00 0.00
Mn2 0.00 0.50 0.00
Mn2 0.00 0.00 0.50
Mn1 0.50 0.50 0.00
Mn1 0.50 0.00 0.50
Mn1 0.00 0.50 0.50
O 0.25 0.25 0.25
O 0.75 0.75 0.75
O 0.75 0.25 0.25
O 0.25 0.75 0.25
O 0.25 0.25 0.75
O 0.75 0.75 0.25
O 0.75 0.25 0.75
O 0.25 0.75 0.75

CELL_PARAMETERS alat
1.0000000000000000000 1.0000000000000000000 - 0.010909629518976782
-0.010909629518976782 1.0000000000000000000 1.0000000000000000000
1.0000000000000000000 - 0.010909629518976782 1.0000000000000000000

Lattice constant = 8.326106103444015 bohr

MnO DMC Time step extrapolation for excited state

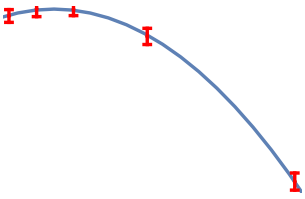
In[ ]:=
```
data = {
    {0.05, -961.969336},
    {0.0250, -961.908226},
    {0.0125, -961.896861},
    {0.00625, -961.896630},
    {0.0015625, -961.899623}
   };
errors = {
    0.003610,
    0.003408,
    0.002606,
    0.003265,
    0.002685};
withError = Transpose[{data[[All, 1]], data[[All, 2]], errors}];
nlm = NonlinearModelFit[data, a x^2 + m x + b, {a, m, b}, {x}, Weights → 1 / errors^2]
nlm[{"BestFit", "ParameterTable"}]
Needs["ErrorBarPlots`"]
Show[ListPlot[data], Plot[nlm[x], {x, 0, 10}],
 ErrorListPlot[withError, PlotStyle → Red],
 PlotRange → {{data[[1]][[1]], data[[5]][[1]]}, {data[[1]][[2]], data[[5]][[2]]}}]
```

Out[ ]= $\text{FittedModel}\left[-961.9 + 0.801096\, x - 43.6445\, x^2\right]$

Out[ ]= $\left\{-961.9 + 0.801096\, x - 43.6445\, x^2,\right.$

	Estimate	Standard Error	t-Statistic	P-Value
a	−43.6445	1.26405	−34.5276	0.000837763
m	0.801096	0.0669826	11.9598	0.00691878
b	−961.9	0.000560463	-1.71626×10^{6}	3.39506×10^{-13}

Out[ ]=



In[ ]:=

MnO DMC Time step extrapolation for ground state

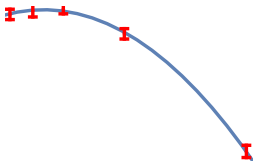
In[ ]:=
```
data = {
    {0.05, -962.117958},
    {0.0250, -962.056149},
    {0.0125, -962.043055},
    {0.00625, -962.043887},
    {0.0015625, -962.045936}
   };
errors = {
    0.003215,
    0.002726,
    0.002579,
    0.003385,
    0.002717
   };
withError = Transpose[{data[[All, 1]], data[[All, 2]], errors}];
nlm = NonlinearModelFit[data, a x^2 + m x + b, {a, m, b}, {x}, Weights → 1 / errors^2]
nlm[{"BestFit", "ParameterTable"}]
Needs["ErrorBarPlots`"]
Show[ListPlot[data], Plot[nlm[x], {x, 0, 10}],
 ErrorListPlot[withError, PlotStyle → Red],
 PlotRange → {{data[[1]][[1]], data[[5]][[1]]}, {data[[1]][[2]], data[[5]][[2]]}}]
```

Out[ ]= $\text{FittedModel}\left[-962.047 + 0.730692\, x - 43.2485\, x^2\right]$

Out[ ]= $\left\{-962.047 + 0.730692\, x - 43.2485\, x^2,\right.$

	Estimate	Standard Error	t-Statistic	P-Value
a	−43.2485	2.15855	−20.0359	0.00248178
m	0.730692	0.115475	6.32772	0.0240767
b	−962.047	0.00105415	−912 624.	1.20071×10^{-12}

Out[ ]=



In[ ]:=

MnO (HSE06 starting Orbitals) Band Gap from DMC

```
In[ ]:= (* optical gap in eV *)
        (-961.9003939533226` + 962.0465547653065`) * 27.2114

Out[ ]= 3.97724

In[ ]:= (* standard error of gap in eV*)
        (√(0.0005604629614023827`^2 + 0.0010541548469112532`^2)) * 27.2114

Out[ ]= 0.0324873
```

Finite size effect correction to DMC gap

DMC band gaps calculated via simple particle-hole excitation of 'homo-lumo' of HSE06 orbital Slater Determinant
Timestep = 0.003125 for all calculations

In[ ]:=
```
data = {
    {0.0625, 4.02445721440163},
    {0.03125, 4.54463033680309},
    {0.015625, 4.28394512479923}
  };
errors = {0.106879815148777,
    0.207994892295513,
    0.630707559700845};
withError = Transpose[{data[[All, 1]], data[[All, 2]], errors}];
nlm = NonlinearModelFit[data, m x + b, {m, b}, {x}, Weights → 1 / errors^2]
nlm[{"BestFit", "ParameterTable"}]
Needs["ErrorBarPlots`"]

(*Show[Plot[nlm[x],{x,data[[3]][[1]],data[[1]][[1]]}],ListPlot[data],
 ErrorListPlot[withError,PlotStyle→Red]]*)
Show[ListPlot[data, PlotRange → {{-0.1, 0.1}, {2, 6}}],
 ErrorListPlot[withError, PlotStyle → Red], Plot[nlm[x], {x, -10, 10}]]
```

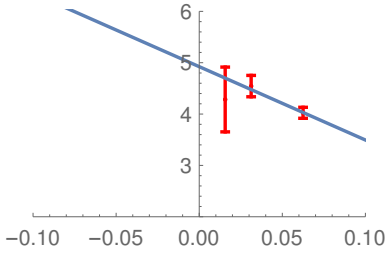
Out[ ]= FittedModel[4.92311 − 14.2828 x]

Out[ ]=

$4.92311 - 14.2828\,x$

	Estimate	Standard Error	t-Statistic	P-Value
m	−14.2828	4.99137	−2.8615	0.214031
b	4.92311	0.283522	17.3641	0.0366225

Out[ ]=



DMC correction via `dmc extrapolation'.

```
(* difference between parameter b and 16-atom band gap*)
4.92311 - 4.02445721440163
```

Out[ ]= 0.898653

```
(* propagated error *)
```

In[ ]:= $\sqrt{0.106879815148777^2 + 0.28352183541302217`^2}$

Out[ ]= 0.302998

Final correction to DMC calculation

```
(* corrected gap *)
3.977240319219197` + 0.8986527855983706`
```

Out[ ]= 4.87589

```
(* propagated error *)
```

In[ ]:= $\sqrt{0.03248727977210656`^2 + 0.3029982277872357`^2}$

Out[ ]= 0.304735

Final correction to VMC calculation

In[ ]:=
```
(* corrected gap *)
5.585 + 0.8986527855983706`
```

Out[ ]= 6.48365

```
(* propagated error *)
```

In[ ]:= $\sqrt{0.05^2 + 0.3029982277872357`^2}$

Out[ ]= 0.307096

Z -> Gamma transition of MnO

Unit cell information (2x2x2 Monkhorst-Pack grid)

(RRKJ ECP for Mn and Shin ECP for O, kinetic energy cutoff = 300 Ry)

4 atom unit cell

lattice constant 8.326099506257146 bohr

CELL_PARAMETERS alat
1.0 0.4945446 0.4945446
0.4945446 1.0 0.4945446
0.4945446 0.4945446 1.0

ATOMIC_POSITIONS crystal
Mn1 0.0000 0.0000 0.0000
Mn2 0.5000 0.5000 0.5000
O 0.2500 0.2500 0.2500
O 0.7500 0.7500 0.7500

G0W0 extrapolation of MnO Z->Gamma transition

extrapolation of G0W0 gap from Abinit using PAW+GW+PBE.
Just to illustrate the extrapolation method is reasonable

The yellow cell (0.8461) represents the extrapolated value while the converged value was (0.8375)
To produce the extrapolated value take the (2x2x2+256_bnds) gap value and correct with the the difference between (4x4x4+28_bnds - 2x2x2+28_bnds)

	2x2x2	4x4x4	8x8x8			
28_bnds						
	10.7048	10.3597	10.3225			
	8.7332	8.1361	8.115			
	1.9716	2.2236	2.2075		0.252	0.2359
256_bnds						
	9.2525	8.8903	8.8408			
	8.6584	8.034	8.0033			
	0.5941	0.8563	0.8375		0.2622	0.2434
	1.3775	1.3673	1.37			
						reconstruct
						0.8461

extrapolation of G0W0 gap from BerkeleyGW using the same technique from the above abinit test.

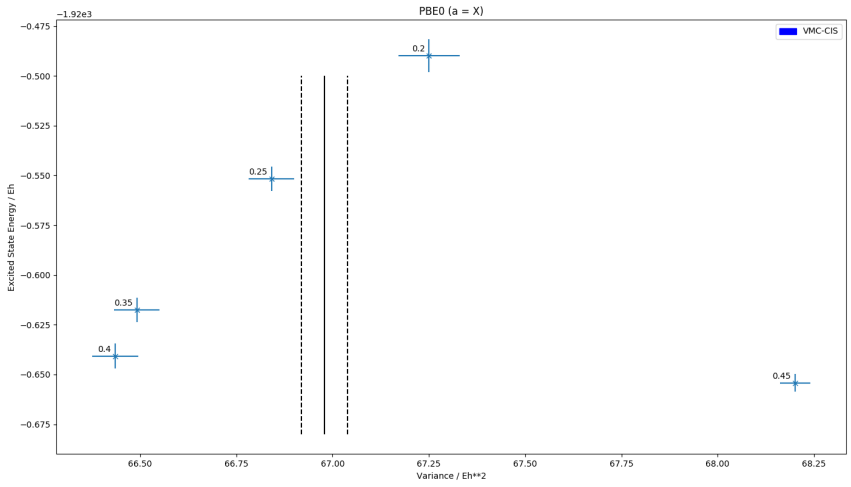
This is the Z->Gamma transition

	2x2x2	4x4x4			
28_bnds					
	17.38453434	17.46862479			
	10.61129595	10.37872174			
	6.773238386	7.089903053			0.316664667
256_bnds					
	14.94513557				
	10.35668814				
	4.588447437				
	2.184790949	7.089903053			4.905112104

This is the Gamma->Gamma transition

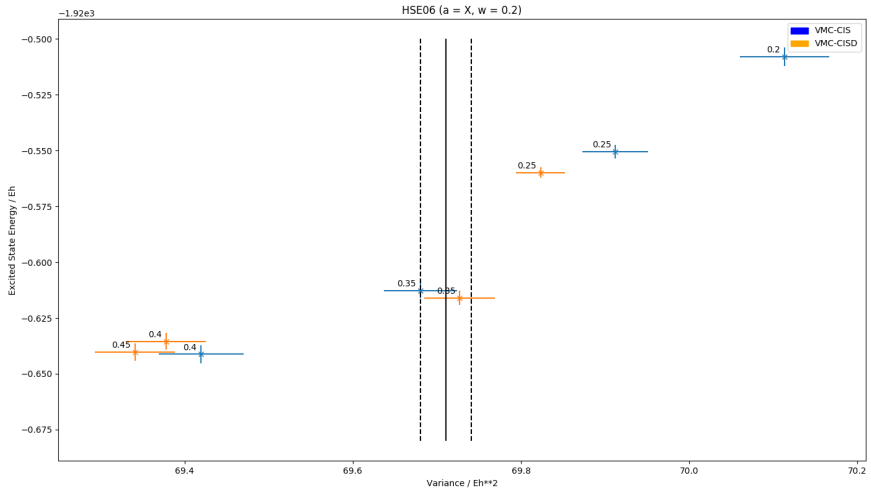
	2x2x2	4x4x4		
28_bnds				
	17.38454131	17.46862479		
	9.794786593	9.550960413		
	7.58975472	7.91766438		0.32790966
256_bnds				
	14.94513557			
	9.665130105			
	5.280005467			5.607915127

Attempts to variance match explicitly by varying the percentage of exact exhange mixture for both PBE0 and HSE06.
The solid black line corresonds to the variance of the ground state and the dash lines are at + /- 1 std_err



In[ ]:=
```
(* Closest variance matched result gap result / eV *)
(-1920.551767 + 1920.800445) * 27.2114
```

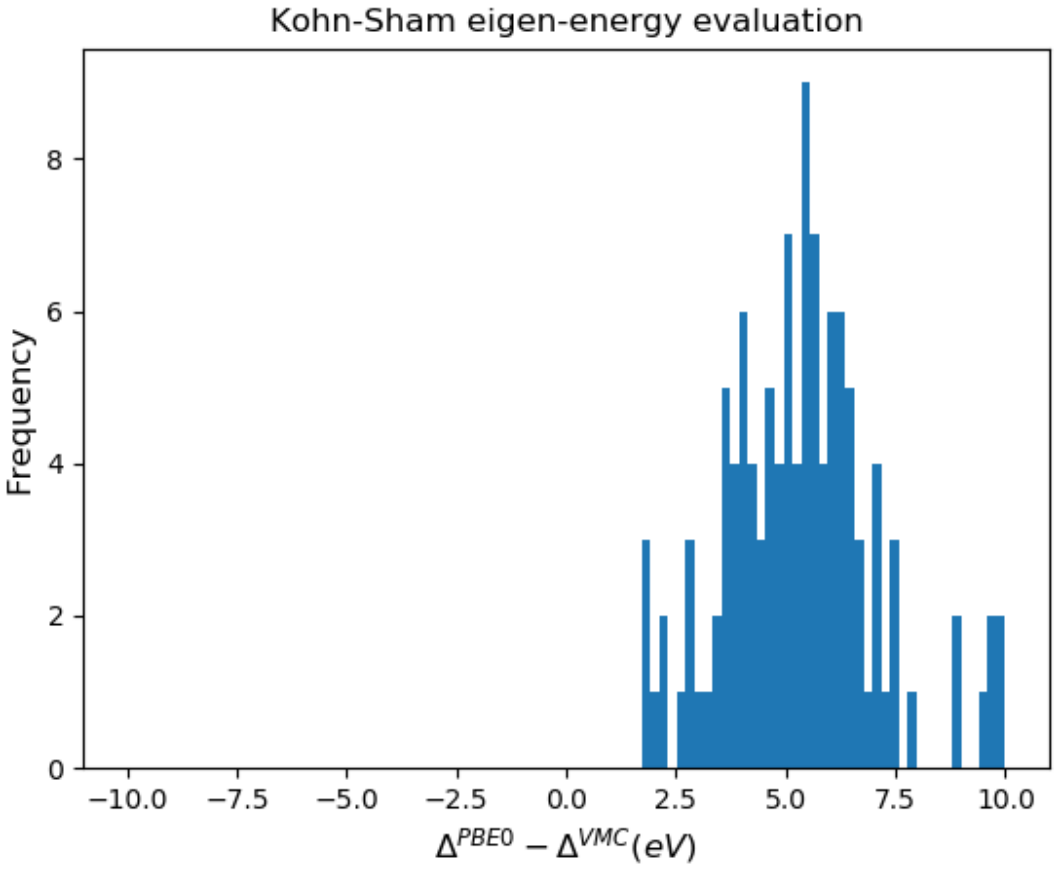
Out[ ]= 6.76688

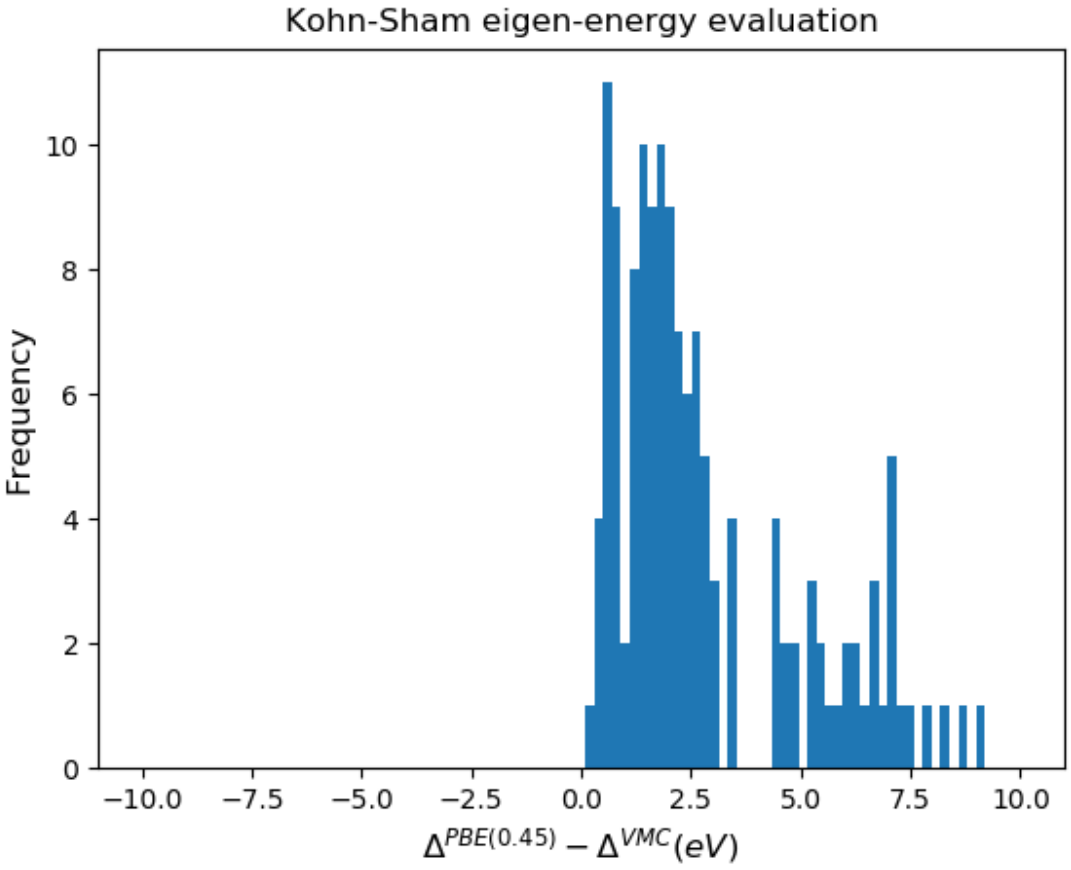


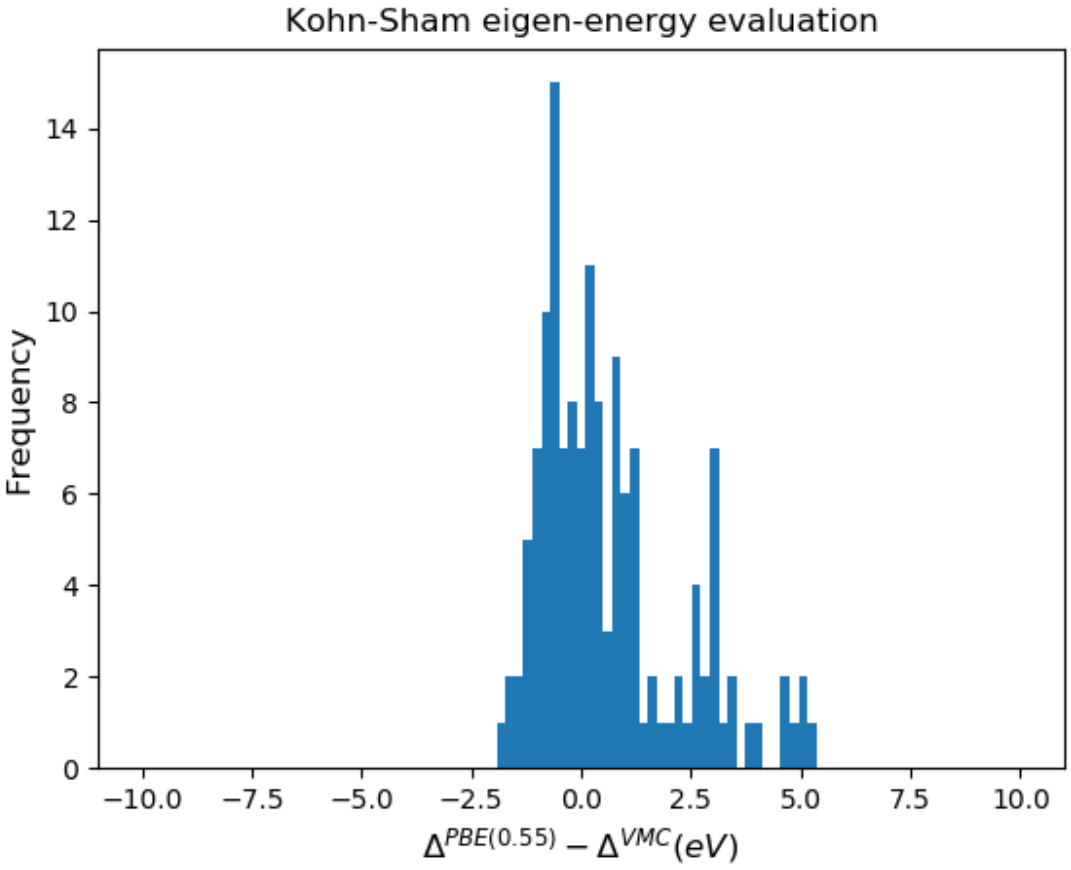
In[ ]:=
```
(* Closest variance matched result gap result / eV *)
(-1920.612763 + 1920.791562) * 27.2114
```

Out[ ]= 4.86537

MnO: exact exchange fraction determination to determine best PBE0 SPO set to use for G_0W_0




Kohn-Sham eigen-energy evaluation
Frequency
0
2
4
6
8
10
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
$\Delta^{PBE(0.45)} - \Delta^{VMC}(eV)$


Kohn-Sham eigen-energy evaluation
Frequency
0
2
4
6
8
10
12
14
−10.0
−7.5
−5.0
−2.5
0.0
2.5
5.0
7.5
10.0
$\Delta^{PBE(0.55)} - \Delta^{VMC}(eV)$