Correlated Methods for Excited States in Periodic Systems

James David McClain

A Dissertation
Presented to the Faculty of Princeton University in Candidacy for the Degree of Doctor of Philosophy

Recommended for Acceptance by the Department of Chemistry
Adviser: Garnet K.-L. Chan

September 2017
Abstract

In this thesis we apply equation-of-motion coupled-cluster theory to obtain the ground-state and excited-state of three-dimensional solids. We show the problems theorists face in applying traditional electronic structure methods to solids and derive the Hamiltonian matrix elements for extended systems. The coupled-cluster due to their size-extensivity and accuracy have long been a promising candidate for use as a benchmark in the ground- and excited-states of solids. After reviewing some basic theory we apply these methods to various periodic systems. We first look at the jellium model with a Wigner-Seitz radius of $r_s = 4$, a model for metallic sodium and compare our findings with state-of-the-art $GW$ theory. Despite the relatively higher scaling of the coupled-cluster methods we were able to achieve modest system sizes of 114 electrons in 485 orbitals, without the use of symmetry. We compute the one-particle coupled-cluster Green’s function and compare the experimental occupied bandwidth for metallic sodium against our calculated one. Moving on to $ab$ initio systems, we perform calculations of optical gaps and band structure of silicon and diamond using the same coupled-cluster methods. Using the k-point equation-of-motion equations provided, we are able to achieve system sizes of 256 electrons in 2,176 orbitals using a $4 \times 4 \times 4$ Monkhorst-Pack sampling of the Brillouin zone and polarized triple-zeta basis. These methods are then presented in the broader context of electronic structure methods provided in the PySCF framework, a software package providing the computational tools for studying the electronic structure of molecular and periodic systems. We conclude with forward directions and preliminary results for removing finite-size effects in correlated methods and pushing towards chemical accuracy through perturbative methods for excited states.
## Contents

Abstract iii

List of Figures vii

List of Tables x

Acknowledgements xi

1 Preface and Introduction 1

2 The Hartree-Fock Approximation and Configuration Interaction 9
   2.1 The Mean-Field Hamiltonian ........................................ 9
   2.2 Moving Beyond a Mean-Field Approach: Configuration Interaction 12

3 Coupled-cluster Theory 16
   3.1 The Coupled-cluster Ansatz ........................................ 17
   3.2 Connection to Perturbation Theory ................................ 20
       3.2.1 Infinite order MBPT ........................................ 20
       3.2.2 Perturbative CCSD .......................................... 23
   3.3 Equation of Motion Theory ......................................... 25
   3.4 EOM-CCSD Equations .............................................. 27
       3.4.1 IP EOM-CCSD Equations .................................... 28
       3.4.2 EA EOM-CCSD Equations .................................... 30
   3.5 Perturbative Triples: EOM-CCSD* ................................. 31

4 Periodic Systems 33
   4.1 The Basics of Periodic Systems ................................. 34
       4.1.1 Unit Cells and Supercells ............................... 34
       4.1.2 Planewaves and k-points ............................... 36
       4.1.3 Fourier Transforms and Useful Identities ............ 41


iv
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>7.2.4 Molecular properties</td>
<td>95</td>
</tr>
<tr>
<td>7.2.5 Relativistic effects</td>
<td>96</td>
</tr>
<tr>
<td>7.2.6 Orbital localizer and result analysis</td>
<td>96</td>
</tr>
<tr>
<td>7.2.7 Extended systems with periodic boundary conditions</td>
<td>97</td>
</tr>
<tr>
<td>7.2.8 General AO integral evaluator and J/K builds</td>
<td>98</td>
</tr>
<tr>
<td>7.2.9 General integral transformations</td>
<td>99</td>
</tr>
<tr>
<td>7.2.10 Density fitting</td>
<td>100</td>
</tr>
<tr>
<td>7.2.11 Custom Hamiltonians</td>
<td>101</td>
</tr>
<tr>
<td>7.2.12 Interfaces to external programs</td>
<td>101</td>
</tr>
<tr>
<td>7.2.13 Numerical tools</td>
<td>102</td>
</tr>
<tr>
<td>7.3 Design and implementation of PySCF</td>
<td>102</td>
</tr>
<tr>
<td>7.3.1 Input language</td>
<td>103</td>
</tr>
<tr>
<td>7.3.2 Advancing interactive computing</td>
<td>104</td>
</tr>
<tr>
<td>7.3.3 Methods as plugins</td>
<td>105</td>
</tr>
<tr>
<td>7.3.4 Seamless MPI functionality</td>
<td>105</td>
</tr>
<tr>
<td>7.4 Conclusions</td>
<td>107</td>
</tr>
<tr>
<td>8 Future Directions</td>
<td>113</td>
</tr>
<tr>
<td>8.1 Finite-size Errors</td>
<td>113</td>
</tr>
<tr>
<td>8.2 Periodic EOM-CCSD*</td>
<td>114</td>
</tr>
<tr>
<td>8.3 Beyond EOM-CCSD*</td>
<td>116</td>
</tr>
<tr>
<td>8.4 Conclusion</td>
<td>118</td>
</tr>
</tbody>
</table>

References

119
List of Figures

1 Spectral functions for the UEG with \( r_s = 4.0 \) using a supercell containing 14 electrons in 19 spatial orbitals. (a) For the \( k = (0,0,0) \) state, the spectral functions exhibits a prominent satellite peak; the HF+GW result has been scaled down by a factor of 1.5. (b) For the highest occupied state at \( k = (2\pi/L,0,0) \), the spectral function exhibits a strong quasiparticle peak with a very weak satellite structure; the satellite region between \(-18 \text{ eV} \) and \(-7 \text{ eV} \) has been magnified by factor of 30. A linewidth broadening of \( \eta = 0.2 \text{ eV} \) is used in all calculations. ........................................................................................................... 61

2 (a) Spectral function of the \( k = (0,0,0) \) state of the 3D UEG with \( r_s = 4.0 \) and 114 electrons in 485 orbitals. The HF+GW result is scaled down by a factor of 1.5 and a linewidth broadening of \( \eta = 0.8 \text{ eV} \) is used in all calculations. (b) Comparison of the spectral functions of the \( k = (0,0,0) \) state in the thermodynamic limit (solid curves) and the 114 electron system (dashed curves) from LDA+GW (blue curves) and LDA+GW+C (green curves). (c) Complete basis set limit quasiparticle energies as a function of wave vector for the 114 electron system (symbols) and quadratic fits (dashed curves). ........................................................................................................... 65

3 Correlation energy per unit cell of diamond calculated with MP2 (left, open symbols) and CCSD (right, filled symbols). Dashed lines indicate the \( N_k \to \infty \) extrapolated values assuming a finite-size scaling of the form \( N_k^{-1/3} \). ........................................................................................................... 87
Cohesive energy of diamond (left) and silicon (right), calculated with the TZVP basis. For each $k$-point mesh, MP2 and CCSD correlation energies are added to a separately converged HF crystal energy. The correlation component of the cohesive energy is extrapolated assuming a finite-size error that scales like $N_k^{-1/3}$. Zero-point corrected experimental results [131] are shown as a thin grey line. Predicted and experimental values are given in Tab. 2.

Equation of state of diamond (left) and silicon (right) with the TZVP basis and a $3 \times 3 \times 3$ sampling of the Brillouin zone.

Band structure of diamond calculated with DFT (PBE), HF, and EOM-CCSD, using the DZVP single-particle basis and a $3 \times 3 \times 3$ $k$-point mesh.

The same as in Fig. 6, but for silicon.

Indirect band gap of diamond (left) and silicon (right) calculated with EOM-CCSD. Dashed lines indicate extrapolation of the form $N_k^{-1/3}$ based on the isotropic $3 \times 3 \times 3$ and $4 \times 4 \times 4$ sampling meshes; grey horizontal lines indicate experimental values. For the TZVP data, the extrapolation predicts for diamond a band gap of 5.37 eV compared to the experimental value of 5.45–5.50 eV [133, 134], and for silicon a band gap of 1.19 eV, compared to the experimental value of 1.17 eV [134, 135].

Example to define a custom exchange-correlation functional for a DFT calculation.

Example to enable the DMRG solver in a CASSCF calculation.

Example to generate localized orbitals and to plot them in Jmol.

Example to access AO integrals.

Example to use a custom Hamiltonian.

Using Python to combine the calculation and data post-processing in one script.
Accessing documentation within the IPython shell. The question mark activates the documentation window in the bottom area. The pop-up menu for code auto-completion is triggered by the <Tab> key.

Example to use plugins in PySCF. The mean-field calculation is decorated by the density fitting approximation, X2C relativistic correction and second order SCF solver.

Comparison of the input script for serial-mode and MPI-mode calculations. Except for the module to import, the MPI parallel mode takes exactly the same input as the serial mode.

The band gap for diamond calculated at the EOM-CCSD level of theory for DZVP and TZVP with varying number of k-points is shown in blue and red, respectively. Also shown is the DZVP calculated at the EOM-CCSD* level of theory, denoted by the pink line.

Performance of EOM-CCSD* against EOM-CCSD for a subset of the GW100 molecular dataset for calculation of (top) ionization potential, (middle) electron affinity, and (bottom) band gaps versus CCSD(T) reference data.
List of Tables

1  Hartree-Fock total energy per unit cell and MP2/CCSD correlation energy per unit cell of diamond in eV. ........................................ 80
2  Lattice constant $a$, bulk modulus $B$, and cohesive energy $\Delta E$ of diamond and silicon using the TZVP single-particle basis. Lattice constant and bulk modulus were obtained from fits to calculations with a $3 \times 3 \times 3$ $k$-point mesh. Crystal energy contributions to the cohesive energy are calculated at the $T = 300$ K lattice constant and extrapolated to $N_k \to \infty$. Values given in parentheses are those calculated with the plane-wave PAW approach as reported in Refs. [8, 113]. All experimental values have been corrected for zero-point vibrational effects [131]. ............... 81
3  Features of the PySCF package as of the 1.3 release. ................. 108
Acknowledgements

First and foremost, I’d like to thank the students of the Chan group. Without them, my experience would have been far less memorable and far less enjoyable. Don’t get me wrong, I did enjoy the work, but when work was slow the group was able to help me through those tedious times. Whether or not we worked specifically on a project together, you were all there to help support me morally and were eager to answer any question I may have had.

I’d like to extend a special thanks to three people in the group. The first would be Thomas J. Watson (not to be confused with the Thomas J. Watson), who would be the first of two people that I would consider my mentors. In my first few years of my postgraduate studies I was running in circles (or maybe just briskly walking), and it wasn’t until I joined up forces with Tom Watson that I actually started to make progress. His enthusiasm for the work was contagious and Tom also instilled in me a deep respect for the coupled-cluster methods and their utility. Tom (and later Tim) were able to answer any of the questions I had, and as my knowledge for the subject grew so did my appreciation of it. Tim would later join our group and help me further these methods for ab initio systems. He was also their to answer questions I had not just about theory, but about the next steps to take after my postgraduate studies. Without their mentorship and friendship I don’t think I would have succeeded as I did in this field. Finally, I would like to thank Qiming Sun. I knew I enjoyed programming before joining the group, but Qiming Sun taught me there was so much else to learn. I thank him for opening my eyes to the world of high-performance-computing.

I’d like to thank my advisers and inspiring teachers over the years. First would be Larry Sullivan, my high school chemistry teacher that instilled a love of chemistry in myself and my brothers, all of whom went on to pursue doctorates in chemistry. Next, I would thank my chemistry adviser Joshua Schrier and mathematics adviser Lynne Butler at Haverford College. Joshua Schrier’s enthusiasm for the field knows no bounds and it’s amazing to see how many truly ingenious ideas he can come up with and the
ways he can test them using simple chemical and computational models. Lynne Butler taught me the importance of being precise (a concept that's importance I did not fully appreciate at the time); when you can’t define a problem then it makes no sense to talk about a solution.

Finally, I thank my adviser Garnet Chan not only for his financial assistance, but also for believing in me and giving me the independence to carry out my studies. I’m always in awe at the vast knowledge he has of a wide range of scientific disciplines. He has an aptness to extract abstract ideas from developments in one project and apply these findings to a seemingly disparate project in another. But even then, he can convince you it was the obvious thing to do by explaining his thought-process in such a cogent and simplistic manner. This ability makes his talks all the more enjoyable to researchers and students, particularly for undergraduates as I found from being a TA in one of his courses. And for all of that I have the utmost respect for him.

I extend my thanks to my committee members Herschel Rabitz, Annabella Selloni, and Robert Cava for their taking part and asking insightful questions in the various milestones throughout my doctorate studies.
Dedicated to my Mom, Dad,

brothers, and cats...
1 Preface and Introduction

While investigation of the electronic structure of solids is something certainly not new, progress in method development from the quantum chemistry community has been slow. While I can’t pinpoint the exact reason it has been slow, there have been a number of factors that I believe has contributed to this. First, there is a lack of open-source software to serve as a foundation for further theory development for periodic systems. Second, there are difficulties that arise both theoretically and computationally when dealing with an infinite system. And finally, current DFT-based methods can do quite well in describing many material properties, begging the question: "why bother?". While the first two points will either be addressed later on in this thesis and/or have been addressed in our open-source software PySCF, the latter point is useful to address here as it lays out the motivation for the remainder for this thesis.

*Why use a Gaussian wavefunction-based method?*

With the widespread use and success of methods such as DFT that rely on the use of density-functionals, there is the question of why we would develop a wavefunction-based method for solids. DFT, and its users, enjoy many benefits which has contributed to its widespread use. One benefit can be attributed to its relatively low-scaling, a cost of
$O(N^3)$ (coming from the diagonalization of the Hamiltonian) where $N$ is the number of basis functions, which allows DFT to be used for large systems. Another is that DFT is a blackbox procedure, where without any tuning from the user and only the selection of a exchange-correlation functional you can expect that your answer can be reasonable. For most applications this answer is acceptable and can be, for very large systems, perhaps the only viable option.

The main problem with DFT is the parameterization and the difficulty of achieving higher accuracy, useful in gauging the exactness of your current solution. With wavefunction methods this latter point can be achieved through using higher level orders of theory, usually at a fairly steep cost, such as expanding your wavefunction ansatz to include higher-order excitations or by including perturbation theory on top of your existing calculation. This is useful, for instance, if you are determining the phase diagram for a system that has very small energy differences between phases (as it usually happens in so-called “exotic” materials). The cost has been the primary hindrance, not to mention that the theory can be difficult too! The other difficulty lies in the parameterization of the exchange-correlation functional of DFT, where the parameters were tuned to fit some properties of various systems. The quality of the fitting and size of the training set for this procedure will determine how well the exchange-correlation functionals generalize to arbitrary systems. For systems in far different environments from which the parameters were trained, there is uncertainty in how DFT will perform. While there certainly is error in the wavefunction method, this error can be more carefully controlled and monitored.

The majority of current software that can run solids uses planewaves as their basis, as opposed to a Gaussians, carrying with it both advantages along with its disadvantages. The major advantage to using a planewave-based code is that reaching the complete basis set limit is much easier than with Gaussians. Reaching the complete basis set limit is easier due to the fact that the planewaves are constructed to be orthogonal whereas the Gaussians need to be made orthogonal when solving the generalized eigenvalue problem for the mean-field hamiltonian. Adding sufficiently diffuse Gaussians to
reach the thermodynamic limit means that the orthogonality between basis functions is harder to achieve, and ultimately you run into linear dependency issues with which the generalized eigenvalue problem can’t be solved. Likewise, when your calculation requires sufficient accuracy, the planewave approach is superior and can achieve high accuracy (think on the order of 1e-8) with much fewer basis functions than Gaussians. For all intermediate regimes of accuracy, however, one can achieve answers much closer to the basis set limit with far fewer Gaussians as compared to planewaves. Less basis functions means reduced memory requirements and increased speed, the latter of which is particularly important for the higher scaling post-Hartree-Fock methods. Extrapolating to the complete basis set limit from this intermediate regime of accuracy is much less straightforward than that of the planewave counterpart.

What codes out there exist for calculations on solids?

Right now we have explained the reasoning behind both a wavefunction-based method and the choice of Gaussians as the basis set. Now we will explore the current software available for doing electronic structure calculations on solids. Because we are interested in doing post-Hartree-Fock calculations, we will only concern ourselves with the software that actually implement the Hartree-Fock and post-Hartree-Fock methods. Of all the software out there, CP2K [1] was the most similar to what we wanted to achieve. CP2K makes use of k-points and uses the localized Gaussian basis functions expanded over planewaves approach [2], allowing for efficient treatment of electron repulsion integrals through the use of Fourier transforms. CP2K also offers the Gaussian and augmented-plane-wave (GAPW) [3], using a combination of both sharp Gaussians and smooth planewaves to represent the density. This is particularly useful for doing all-electron calculations, where traditionally a large number of planewaves were needed to represent sharp nuclear potential features in the absence of pseudopotentials. While CP2K was a promising candidate, the choice of Fortran programming language would hinder quick development and would have a steeper learning curve in learning its functionalities.
Besides CP2K, there was a long list of other software that were thrown out as potential candidates for serving as a foundation for our Hartree-Fock and post-Hartree-Fock methods. This was due to some combination of either being written in languages that would prohibit quick development, generally Fortran, and/or not having the functionality for the basis sets we wanted. Quantum Espresso [4] is a planewave code allowing for electronic structure calculations and simulations on solids, traditionally built for DFT exchange but later allowing for Hartree-Fock exchange. The Vienna Ab initio Simulation Package (VASP) [5–7] allows for Gaussian basis sets but is centered around the use of more difficult to use Projector augmented-wave (PAW) approach for core electrons, cumbersome for method development in Hartree-Fock and post Hartree-Fock methods. While many of these packages have some post Hartree-Fock methods such as MP2 implemented, it is notable that VASP has the ability to also perform the ground state coupled-cluster methods CCSD and CCSD(T) [8]. And lastly is the CRYSCOR [9] suite for programming that was not considered due to it being closed source.

The requirement of being open-source, written in a high-level language for quick development, and having both Hartree-Fock and Gaussian basis sets was not met by any of the available software on the market. For that reason we decided to develop our own code within the Python-based Simulation of Chemistry Framework (PySCF) suite [10]. While almost all of the code is written in Python, bottlenecks within the code were written in C for more efficient implementations, allowing it to compete in terms of timings with other pure Fortran/C implementations of electronic structure packages.

**What difficulties arise in periodic systems?**

While the depth of literature on electronic structure is daunting, the application and development of these methods for solids has really only taken off in the recent years. The primary difficulty in solid calculations is the long slow-decaying Coulomb potential. In a molecular calculation, this is of course not a problem as your system is generally well-localized spatially; two electron densities sufficiently far away from the system should have negligible interaction. In a periodic system, however, you are calculating the potential in a unit cell due to an infinite number of replicas of this
unit cell. The density is periodic with the system and one does not have this decay of
the one-electron density matrix, leading to a sum of Coulomb interactions for constant
densities over all periodic images and a divergent interaction. So not only does your
calculation generally become expensive and difficult from a computational perspective,
but this divergence also raises some concerns related to the theoretical foundations of
how one treats the periodic system.

The primary point to be made here is that there subtleties in doing electronic
structure theory in solids that one won’t find in the molecular case. This makes the
problem of doing electronic structure in solids a more difficult endeavor; while the
hurdles to the theory can be overcome, one is still ultimately left with the computational
burden of these calculations. All of this has hindered the development of the field of
electronic structure theory for periodic systems. But in this difficulty one simply has
to take a new approach to the problem. In molecular calculations one implements
symmetry as almost an afterthought. Here the symmetry found in the translationally
invariant systems here will have to be built in from the very beginning to offset the
high computational demands of both the system we wish to explore and the methods
with which to explore it with. Without the use of periodic boundary conditions and
the symmetry this imposes on the problem, these calculations presented here would not
be feasible.

Now that we have a feeling for the underlying motivation and history for the project,
we can move on to the thesis and what was actually accomplished. The following thesis
will not be a comprehensive review of everything attempted during my postgraduate
studies, but will be more tailored towards my contributions to the scientific literature.
In addition I will include some (unpublished) insights and derivations that I found
interesting and/or might be helpful to the reader doing work in the field. Again,
electronic structure theory for solids from a chemistry point-of-view has only recently
been developed recently but has roots deep inside the solid-state physics community.
The primary goal of my work was to develop correlated methods for periodic systems. Initially, this was done using a parameterized Hamiltonians and later extended to jellium systems. Our group had been planning to write an \textit{ab initio} Hartree-Fock and DFT for solids and began work on this while I was working with the jellium systems. While I was able to extend some of my knowledge gained from calculations on model Hamiltonians, my primary contribution was to the coupled-cluster methods for ground and excited states, in particular the coupled-cluster singles and doubles (CCSD) and equation-of-motion (EOM-) CCSD. I should note that this was not the first time either of these have been implemented, and note the reader to Hirata et al.’s [11, 12] first use of CCSD in a periodic setting for one-dimensional conjugated polymers in a small basis. Finally Booth, Gruneis and others [8] showed the CCSD and perturbative triples CCSD(T) method could be extended to three-dimensional systems with reasonably sized basis sets.

The primary novelty for this work was the development of EOM-CCSD for three-dimensional systems with periodic boundary conditions. The first instance of using EOM-CCSD for extended systems comes from Katagiri [13] in his work on polyethylene using a small basis. There are a few things that make our work novel: first, that our method works for three-dimensional systems where a special treatment of the Coulomb operator is necessary in order to get a well-defined result and second, that we employ k-points for our EOM-CCSD so that information on the momentum is not lost, leading to an increase in computational efficiency. The former point is important since the Coulomb summation is absolutely convergent for sums over periodic images in one- and two-dimensional systems but is conditionally convergent for those in three-dimensional systems. The use of k-points is necessary if one wants to employ reasonably sized basis sets in your calculation.

Generally if one wants a band gap for a solid one would turn to DFT and \textit{GW} methods. It is well known that Koopman’s theorem does not apply for DFT and so the difference between the HOMO and LUMO levels should not correspond to a band gap, although in practice this approximation is used. This approximation leads to band
gaps that are systematically lower than the experimental band gap. To improve on this, one can turn to the $GW$ methods [14, 15]. In this context, the $GW$ method is usually viewed as a perturbation to some mean-field solution, generally from DFT, that introduces a screened Coulomb operator in place of the bare Coulomb operator. The strength of the screening is mediated by the dielectric function and hence polarizability of the material. For metals and small band-gap semiconductors the polarizability can be quite high leading to a much smaller Coulomb interaction compared to that used in the bare model. The $GW$ methods have been widely used for metallic systems and small band-gap semiconductors, where there is a large screening and Hartree-Fock and its simple perturbation methods will fail. And for even larger band gap systems, the $GW$ method perform quite well and greatly reduces the degree to which the DFT band gap is underestimated. Still, however, there is a tendency to underestimate the band gap on the order of a few tenths of an eV [16, 17].

While the $GW$ method performs well in correcting the DFT energies and producing band gaps closer to the experimental values, there are some subtleties and deficiencies to this method. As the $GW$ is usually done in a perturbative manner (usually called one-shot $G_0W_0$) rather than self-consistently, which again has some arbitrariness, there is the issue of the dependency on the starting guess. One finds that the discrepancy can be quite large for starting with methods such as Hartree-Fock and a DFT local density approximation (LDA). For methods such as coupled-cluster, while one can start with an arbitrary reference one finds that the dependence on the initial guess is small. The other issue is that improving upon the $GW$ solution can be quite difficult; extensions to $GW$ such as the $GW$ plus cumulant theory ($GW + C$) can help correct the deficiencies in the $GW$ satellite structure but generally leave the location of the quasiparticle peaks (QP), and hence band gap, relatively untouched [18–20]. This difficulty also means that getting an estimation of the accuracy of the calculation is not straightforward, another situation that is not helped by the lack of available methods for accurate excited state calculations in solids.
Our use of EOM-CCSD in solids was originally used to calculate the satellite structure of the three-dimensional uniform electron gas, a simple periodic system consisting of a uniform electron density that can be used as a model for metallic systems such as sodium. The structure of satellites and other excited state properties can be obtained from the spectral function, related to the imaginary part of the Green’s function. Properties elucidated by looking at the spectral function include excitation energies, given by the position of the peaks, along with the strength and lifetime of the peak, related to the peak height and width of the peaks. From a mean-field Hartree-Fock point of view, the peaks are delta functions at the mean-field eigenvalues. By introducing correlation, these delta functions are broadened and new peaks are introduced by allowing for collective excitations within the system. The $GW$ method has been known to predict a single satellite peak instead of multiple peaks with decreasing weight along with an overbinding of the satellite peaks [21]. The $GW + C$ method has been known to correct this through the introduction of higher order diagrams, termed vertex corrections. Because the spectral function was the primary point of interest, a Green’s function approach was taken for coupled-cluster as outlined by Nooijen [22, 23]. Our work on the electron gas represented the first implementation of the CCSD Green’s function approach, and while at that time we had not included symmetry within our ACESIII [24] implementation we were still able to systems of modest size, including sizes up to 114 electrons in 485 spatial orbitals.
2 The Hartree-Fock Approximation and Configuration Interaction

We will begin by exploring the Hartree-Fock approximation as a mean-field method and Configuration Interaction (CI) as a way to improve the mean-field theory. Almost all the calculations we have done have started with the Hartree-Fock approximation while the post-Hartree-Fock methods we employ all were built as ways to improve the deficiencies in the CI method. As it stands, the CI method is the easiest way to add in correlation, even though it may not be the best at adding correlation. For this, we will list some of the deficiencies of the method, particularly as it applies to periodic systems, and why other alternative methods needed to be explored for calculations in solids.

2.1 The Mean-Field Hamiltonian

At the heart of what we do is solve the Schrodinger equation

\[ \hat{H} |\Psi\rangle = E |\Psi\rangle \]  \hspace{1cm} (1)
given a Hamiltonian operator $\hat{H}$. The Hamiltonian in the case of electronic structure theory is the many-body Hamiltonian

$$\hat{H} |\Psi\rangle = -\frac{\hbar^2}{2m} \sum_i \nabla_i^2 - \sum_{i,A} \frac{Z_A e}{|r_i - r_A|} + \frac{1}{2} \sum_{i \neq j} \frac{e^2}{|r_i - r_j|} - \frac{\hbar^2}{2M_A} \sum_A \nabla_A^2 + \sum_{A,B} \frac{Z_A Z_B}{|r_A - r_B|}$$

(2)

where the variables $m_i$ and $M_A$ represent the mass of the electron $i$ and nucleus $A$, while $r_i$ and $r_A$ represent their coordinates, and lastly $\hbar$ and $e$ are the reduced Planck’s constant and elementary charge of the electron. The first and fourth term represent the kinetic energy of the electron and nucleus, respectively. The remaining second, third, and fifth term represent the electron-nuclear attraction, the electron-electron repulsion, and finally the nuclear-nuclear repulsion energy. To make this problem more tractable, the Born Oppenheimer approximation is used, where the combined electronic and nuclear wavefunction $|\Psi(\mathbf{r}, \mathbf{R})\rangle$ is written as a product of a single electronic and nuclear wavefunction $|\Psi(\mathbf{r}; \mathbf{R})\rangle |\Psi(\mathbf{R})\rangle$ (which is exact in the absence of the electron-nuclear attraction). The Born-Oppenheimer approximation makes use of the fact that the coupling terms neglected in this approximation include contributions that scale as $m_i/M_A \approx 1/10,000$ to the electronic energy, and thus for the purposes of (most) electronic structure calculations can be ignored. Thus we can ignore the fourth nuclear kinetic term in the above and treat the electrons as moving in a field of fixed nuclei and the final nuclear-nuclear term is based on these fixed coordinates.

Thus we’ve arrived at a more tractable set of equations to solve by invoking the Born-Oppenheimer approximation and arriving at the “clamped-nuclei” Schrodinger equation. Now we can solve an electronic Schrodinger equation under the assumption of fixed nuclei centers. The easiest approximation to make for a system of $N$ electrons is that the wavefunction can be factorized into a single antisymmetrized product, a so-called Slater determinant, of single particle spin orbitals $\phi_i(\mathbf{x})$ where the label $i$ encompasses both a spatial orbital part $\psi_i(\mathbf{r})$ and spin $\sigma(\omega) \in \{\alpha, \beta\}$:
\[ |\Psi_0\rangle = |\phi_1(x_1)\phi_2(x_2)\cdots\phi_N(x_N)\rangle. \] (3)

The goal is to find a good candidate set of \( \phi_i(x) \) that will be a starting point for subsequent improvement in higher scaling theories. Moving to a non-mean-field Hamiltonian generally means that these single particle orbitals are no longer eigenfunctions of the new Hamiltonian and our new eigenfunctions will be a linear combination of Slater determinants comprised of these single particle orbitals. There arises the question of whether this notion of a single Slater determinant will help us in finding the true many-body wavefunction. It does, and will rely on the fact that for a complete basis of single particles we can write any many-body wavefunction as a linear combination of a complete set of single determinants. In general our basis will generally be far from complete, and so we will not get the “true” many-body wavefunction. However, we can use our physical intuition to get a good approximation for the many-body wavefunction by systematically improving our basis set and stopping at a desired accuracy.

The Hartree-Fock wavefunction \( |\Psi_0\rangle \) is the single Slater determinant made from these single particle orbitals that gives the lowest energy to the many-body Hamiltonian. The equations are found by variationally minimizing the expectation value of the many-body Hamiltonian under the constraint the single particle orbitals \( \phi_i(x) \) are orthogonal. In this, one solves for the single particle orbitals as eigenfunctions of a mean-field, one-body, Hamiltonian operator \( f(x_1) \),

\[
\begin{align*}
  h(x_1)\phi_i(x_1) + \sum_{j \neq i} \left[ \int dx_2 |\phi_j(x_2)|^2 r_{ij}^{-1} \right] \phi_i(x_1) \\
  - \sum_{j \neq i} \left[ \int dx_2 \phi_j^*(x_2)\phi_i(x_2)r_{ij}^{-1} \right] \phi_j(x_1) &= \epsilon_i \phi_i(x_1)
\end{align*}
\] (4)

giving rise to eigenvalues \( \epsilon_1, \epsilon_2, \ldots, \epsilon_N \) for our \( N \)-electron system, where we have defined the one-electron operator \( h(x_1) \) to include contributions from the kinetic energy and electron-nuclear attraction terms.

One can immediately see that the one-electron operator along with the Coulomb and exchange summations, all depend on a single electron coordinate, \( x_1 \), and the collection
of these terms is grouped to form the Fock operator, \( f(x_i) \). While only \( N \) terms are used for the ground state energy, typically a calculation is done over \( 2K > N \) spin orbitals forming a set of both occupied orbitals, where orbital index \( i \leq N \), and virtual orbitals, where \( i > N \). The \( K \) here represents the number of spatial orbitals used in the calculation while the factor of two corresponds to the ability to create these orbitals for both \( \alpha \) and \( \beta \) spin electrons. Both the occupied and virtual orbital energies do have a physical meaning however, which is the (approximate) ionization potential and electron affinities. While the virtual orbitals are not used in the Hartree-Fock ground state energy, they are essential for obtaining more accurate ground state energies when moving beyond the mean-field theory; the reason for this will become clear shortly. For now, however, we can say that this set of orbitals will help us in forming our complete basis of Slater determinants with which we can write the true many-body wavefunction.

\[ \text{(5)} \]

### 2.2 Moving Beyond a Mean-Field Approach: Configuration Interaction

From Hartree-Fock, one can obtain a set of \( 2K > N \) spin orbitals. While only \( N \) of these orbitals are used for the Hartree-Fock determinant the remaining \( 2K - N \) orbitals can help to create additional determinants. Perhaps the simplest way to do this in a hierarchical fashion is in a configuration interaction (CI) manner, where a complete basis of determinants is formed from all singly, doubly, and higher order excitations out of the Hartree-Fock wavefunction,

\[ |\Psi\rangle = |\Psi_0\rangle + \sum_{ia} C_{i}^{a} |\Psi_{i}^{a}\rangle + \sum_{i<j, a<b} C_{ij}^{ab} |\Psi_{ij}^{ab}\rangle + ... \]

where the notation \( |\Psi_{i}^{a}\rangle \) refers to the singly excited determinant formed from the orbital \( \phi_i \) being replaced by the orbital \( \phi_a \) in the Hartree-Fock wavefunction. Because there are \( 2K \) orbitals to choose from and \( N \) orbitals for each Slater determinant, we have \( \binom{2K}{N} \) different combinations of possible Slater determinants. Here and for the rest of the paper, we will exclusively refer to orbital indices \( i, j, k, ... \) to be in the set of occupied molecular orbitals while \( a, b, c, ... \) will refer to the set of unoccupied (virtual) molecular
orbitals. Note the summation in the CI procedure above is not over the full set of \( i, j, a, b \), due to double counting the antisymmetrized product of single particle orbitals. For example, for a doubly excited determinant we have

\[
|\Psi_{ij}^{ab}\rangle = -|\Psi_{ji}^{ab}\rangle = -|\Psi_{ij}^{ba}\rangle = |\Psi_{ji}^{ba}\rangle
\]  

(6)

from antisymmetrization and thus only the terms where \( a < b \) and \( i < j \) are needed to be included in the above CI expansion. It is worthwhile to note that the inclusion of these additional determinants wouldn’t affect the energy of the \textit{mean-field} Hamiltonian; it will however affect the energy of the \textit{many-body} Hamiltonian. This is easy to see as there could be no non-zero coupling between elements of the mean-field Hamiltonian and any excited determinants.

The CI case reduces to a fairly easy, albeit expensive, linear algebra equation involving the hermitian Hamiltonian

\[
H = \begin{pmatrix}
\langle\Psi_0|H|\Psi_0\rangle & 0 & \langle\Psi_0|H|D\rangle & 0 & 0 & \ldots \\
0 & \langle S|H|S\rangle & \langle S|H|D\rangle & \langle S|H|T\rangle & 0 & \ldots \\
\langle D|H|\Psi_0\rangle & \langle D|H|S\rangle & \langle D|H|D\rangle & \langle D|H|T\rangle & \langle D|H|Q\rangle & \ldots \\
0 & \langle T|H|S\rangle & \langle T|H|D\rangle & \langle T|H|T\rangle & \langle T|H|Q\rangle & \ldots \\
0 & 0 & \langle Q|H|D\rangle & \langle Q|H|T\rangle & \langle Q|H|Q\rangle & \ldots \\
\vdots & \vdots & \vdots & \vdots & \vdots & \ddots
\end{pmatrix}
\]  

(7)

where \(|S\rangle, |D\rangle, |T\rangle, |Q\rangle, \ldots\) represent the single, doubly, triply, and higher order excitations. The Hamiltonian is relatively sparse as we have made use of the fact the Hamiltonian is at most a two body interaction involving four orbitals and can only connect elements that differ by at most four elements. Thus, for example, the corresponding element \( \langle\Psi_0|H|T\rangle \) between the Hartree-Fock determinant and the triply (or higher) excited determinant should be zero. Additionally, the term \( \langle\Psi_0|H|S\rangle \) and its hermitian conjugate are zero as a consequence of Brillouin’s Theorem, which states the singly excited determinants do not couple with the Hartree-Fock ground state solution. This can be seen as a stability condition, where the Fock energy will be unchanged if we
perturb the wavefunction through any singly excited determinant. Inclusion of higher excited states will also leave the answer unchanged as the one-body Fock operator can not couple the Hartree-Fock ground state with these determinants. That does not mean that singly excited states should not be included however, since they lower the energy through coupling with doubly and triply excited states and so will affect the ground state energy, albeit in an indirect way.

While the CI has its good qualities, it has many downfalls. The computational ease in which it can be coded is perhaps one of the best things about this method, requiring only knowing the matrix elements to be plugged into a linear eigensolver. For benchmark calculations, the exact answer in the single particle basis is sometimes sought, and for this the ease of CI allows an easy way to achieve this exact answer. This CI is termed full CI and includes all orders of excitations, for which there are again $\binom{2K}{N}$ of them, and this gives rise to an exponential cost to the algorithm. Typically the full CI can be performed on systems sizes on the order of 10 electrons with a modest sized basis. Even CISD calculations are fairly expensive, scaling as $O(n_{\text{vir}}^4 n_{\text{occ}}^2)$ which for the quality of the calculations is fairly prohibitive. In addition, the CI is variational, meaning it will give an upper bound to the true energy and can be beneficial in that you know the sign of your error.

Variationality in itself isn’t all that useful if you don’t know the magnitude of the error bar, and typically the error for the CISD is much higher compared to other methods with similar scalings. The CISD and other truncated CI methods (including any excitations less than that included in full CI) suffer from additional problems that can be avoided by using a different method to treat electron correlation. For example, the CISD and truncated CI methods are not size consistent, meaning that the energy of two non-interacting monomers is not the same as twice that of the individual monomer. To see this, we can see that CISD is exact in the single particle basis for an individual hydrogen molecule, but for two non-interacting hydrogen molecules separated infinitely far apart we would need up to quadruple excitations, i.e. CISDTQ, to allow for double excitations on each of the hydrogen molecules. Thus we get a significantly worse quality
of energy for the non-interacting monomers, and because of this dissociation energies are generally poor for methods that aren’t size consistent.

There is also the related notion of a method being size extensive, where the energy grows linearly with respect to the number of particles. It is textbook knowledge that for configuration interaction doubles (CID) the correlation energy per monomer tends to zero with increasing number of monomers [25]. For condensed matter systems a method that is not size extensive can be particularly problematic as it predicts the correlation energy per unit cell would be zero in the limit of an infinite number of unit cells. There are ways to combat this and produce a method that is approximately size extensive, such as the Davidson’s correction [26], but these are only approximate corrections. As CISD is not size extensive, as one moves to larger and larger systems the quality of the energy calculated is degraded.

The issues of size consistency and extensivity have been some of the main reasons that CISD has not been used widely in practice. Luckily, there are alternative methods that fix these issues while keeping roughly the same scaling as the CI methods. Perhaps the most widely used are the coupled-cluster methods, namely the coupled-cluster singles and doubles, CCSD, and it’s perturbative CCSD(T) correction. CCSD, its excited state variants, and their application to solids will be the focus for the remaining of this thesis.
3  Coupled-cluster Theory

We found in the last chapter that a truncated CI method would not fair well in solids due to the issue of size-extensivity. The coupled-cluster methods are known to correct this while fixing the problem of size consistency. In addition, taking CISD and coupled-cluster singles and doubles (CCSD), we find that they have the same asymptotic $O(N^6)$ scaling. All of this coupled with the fact that CCSD generally just has better energies when compared to full CI [27] has made CCSD generally a more popular method. While CCSD was originally developed for ground state properties, the equation-of-motion (EOM) theory extends this to excited states where again it has become the method of choice for dealing with single-reference systems. The section here is not meant to be exhaustive and the reader is expect to have some familiarity with perturbation theory. For more extensive introductions to these methods the reader is directed to Shavitt and Bartlett’s book [28].

Here, we will first discuss the coupled-cluster ansatz and the truncation of the exponential ansatz. We will then show how the coupled-cluster theory is intimately related to perturbation theory by summing over a subset of the perturbation diagrams up to infinite order for a given truncation of the exponential operator. This will then be used in the analysis of the CCSD diagrams and lead to the perturbative triples correction
for CCSD, termed CCSD(T). Finally, we will end this section with a discussion of the extension of coupled-cluster and their perturbative corrections to excited states.

### 3.1 The Coupled-cluster Ansatz

Coupled-cluster theory defines an exponential ansatz of the wavefunction

$$|\Psi\rangle = e^T |\Psi_0\rangle ,$$

where $T$ are some excitation operators, written in second quantization as

$$T = T_1 + T_2 + T_3 + ...$$

$$\sum_{ia} t_i^a a_i^\dagger a_i + \left(\frac{1}{2!}\right)^2 \sum_{ijab} t_{ijab} a_i^\dagger a_b^\dagger a_j^\dagger a_i + \left(\frac{1}{3!}\right)^2 \sum_{ijkabc} t_{ijkabc} a_i^\dagger a_b^\dagger a_c^\dagger a_k a_j^\dagger a_i + ...,$$

that are typically cut off at some order of excitation. The normalization factor $(m!)^2$ is chosen since any permutation of $m$-hole and $m$-particle operators only adds redundant terms that are equivalent up to a minus sign due to the use of antisymmetrized operators, i.e. for the $T_2$ operator we have four excitation operators related through the symmetry $t_{ij} = -t_{ji} = -t_{bj} = t_{bj}$. Truncating at singles and doubles excitations, we arrive at the CCSD, where the amplitudes are defined as $T = T_1 + T_2$. Other less common methods are cutting off at triples excitation, for CCSDT, and quadruples, for CCSDTQ. In addition to this, one can also use a noniterative perturbative corrections for coupled-cluster methods. The most famous of these is the noniterative perturbative triples correction for CCSD, termed CCSD(T), that is regarded as the "gold standard" of quantum chemistry and can arrive at chemical accuracy with only a slightly higher cost compared to CCSD.

Applying the many-body Hamiltonian to this ansatz followed by $e^{-T}$, we have

$$e^{-T} H e^T |\Psi_0\rangle = E |\Psi_0\rangle .$$

One can define a "new" Hamiltonian $\hat{H} \equiv e^{-T} H e^T$. This Hamiltonian differs from the
many-body Hamiltonian only by a similarity transform, which should only change the matrix elements and eigenvectors but keep the eigenvalues unchanged. Thus this procedure should obtain the same eigenvalues as the full CI procedure, even if approximate \( T \) amplitudes, are used as long as the transformed Hamiltonian \( \tilde{H} \) is diagonalized exactly in the full set of determinants. Of course diagonalizing in the full set of determinants is not desirable, and so only an approximate diagonalization occurs. This happens in the “coupled-cluster equation”, where one sets \( \langle \Psi_{ijk}^{abc} | \tilde{H} | \Psi_0 \rangle = 0 \) up to some cut-off excitation, which will be seen later. Before diving further into the details about the structure of this new Hamiltonian it would be wise to arrive at an energy expression amenable for computer implementation. I will not get too bogged down in the details, which can be found elsewhere [28, 29], but will only explain what is necessary to further later discussion while helping the reader brush up on any forgotten material.

The derivation of the ground state energy can be obtained by using the Baker-Campbell-Hausdorff formula

\[
e^{-T} H e^T = H + [H, T] + \frac{1}{2!} [[H, T], T] + \frac{1}{3!} [[[H, T], T], T] + \frac{1}{4!} [[[H, T], T], [T], T] + \ldots
\]

(11)

Realizing that again because the Hamiltonian contains at most two-body terms and thus must contract with at most four indices coming from at most four different operators. So only the Hamiltonian up to the fourth commutation can be non-zero, beyond that the commutation will destroy any remaining terms since these will contain operators that do not contract with the Hamiltonian and will be destroyed in the commutation.

Thus we look at all terms that are \textit{connected} with the Hamiltonian, i.e. those that have at least one shared orbital index with the Hamiltonian. While the terms that survive involve the Hamiltonian contracting with at most four different operators, that does not mean that these terms will make it into the final energy expression. For the energy expression we must have the \( T \)'s fully contracted with the Hamiltonian and is given by

\[
E_{CCSD} = \sum_{ia} f_{ia} t_{ia}^a + \frac{1}{4} \sum_{ijab} \langle ij || ab \rangle \left( t_{ij}^{ab} + 2t_{ia}^{ab} t_{ija}^{b} \right).
\]

(12)
Since we have a Hamiltonian that is at most four indices, the only terms that survive are those with at most four indices. This means that the Fock matrix can contract only with $t^a_i$ and the molecular integrals can only contract with either a four-index $t^{ab}_{ij}$ amplitude or product of $t^{ab}_{ij}$ amplitudes. If the amplitudes did not fully contract with the Hamiltonian, then there would be an extraneous hole or particle operator acting on our $|\Psi_0\rangle$, giving us a new state $|\Psi'_0\rangle$. This would then given us zero overlap with $\langle \Psi_0 |$ when closing the expression in equation 10.

The energy in expression 12 is exact, in the full CI sense, given the true amplitudes $t^a_i$ and $t^{ab}_{ij}$. The only remaining difficulty lies in determining these amplitudes. To do this, we note that in equation 10 we need to satisfy

$$\langle \Psi_{ijk}^{abc} | \bar{H} | \Psi_0 \rangle = 0$$

for all excitation operators up to order $N$ for an exact energy. For these equations, it is not necessary for the Hamiltonian to be fully contracted with the $T$ amplitudes, but it is necessary for the Hamiltonian to be fully contracted with both the $T$ amplitudes and the left acting excited state. So although the $t^a_i$ and $t^{ab}_{ij}$ are the only necessary quantities needed for an exact energy, these amplitudes depend on the higher order amplitudes through the set of non-linear equations defined by the constraint above.

For the CCSD method, only the first two set of equations

$$\langle \Psi_{ij}^a | \bar{H} | \Psi_0 \rangle = 0$$
$$\langle \Psi_{ij}^{ab} | \bar{H} | \Psi_0 \rangle = 0$$

are solved. One can see that the first equation defines a set of $n_{occ}n_{vir}$ equations and the second $n_{occ}^2n_{vir}^2$, thus defining a solvable system for our $t^a_i$ and $t^{ab}_{ij}$ amplitudes.

Using the CCSD equations above, one can define a set of update equations for both the $T_1$ and $T_2$ amplitudes. The equations for the singles and doubles amplitudes both depend on the full set of $T_1$ and $T_2$ amplitudes. The CCSD (nonlinear) equations must be solved with a nonlinear solver, usually in an iterative fashion akin to the Jacobi method for linear equations. In this, we can define a new set of equations.
\[
\langle \Psi_i^a | \hat{H} | \Psi_0 \rangle' = (f_{ii} - f_{aa}) t_{i}^a
\]
\[
\langle \Psi_{ij}^{ab} | \hat{H} | \Psi_0 \rangle' = (f_{ii} + f_{jj} - f_{aa} - f_{bb}) t_{ij}^{ab}
\]

where the prime means that we have added the right hand terms in the equation above into each of the original CCSD equations. Solving these modified equations is totally equivalent to solving the original CCSD equations in 14. In the case of a Hartree-Fock reference, we are exactly cancelling the terms involving the Fock matrix element contracting with the \( T_1 \) and \( T_2 \) amplitudes. For non Hartree-Fock references, we are merely subtracting the contraction of the off-diagonal elements with the \( T_1 \) and \( T_2 \) amplitudes remain on the left hand side of the equations above. Written in this form, it is clear we can plug in a previous guess for \( T_1 \) and \( T_2 \) into the equations above and solve for new updates on the right hand side, iterating until convergence.

One of course needs to start with an initial guess, and one way in which to do this is to start with the MP2 amplitudes, which one obtains from equation 15 by assuming the amplitudes are initialized to zero, and thus the first updates will be

\[
t_{i}^{a(1)} = \frac{f_{ia}}{f_{ii} - f_{aa}}
\]
\[
t_{ij}^{ab(1)} = \frac{\langle ab | ij \rangle}{f_{ii} + f_{jj} - f_{aa} - f_{bb}},
\]

where the superscripts represent they are in a sense the first-order approximate amplitudes. By plugging in these amplitudes into the coupled-cluster energy equation one obtains the usual MP2 energy. The fact that this occurs is not a coincidence, as we will show in the next section.

### 3.2 Connection to Perturbation Theory

#### 3.2.1 Infinite order MBPT

Here it is assumed the reader has some knowledge of Many Body Perturbation Theory (MBPT). For a short introduction I lead the reader to an article by Bartlett [27], and for a slightly longer introduction I lead the reader to a book by Bartlett and Shavitt [28]. This connection to MP2 can be furthered by taking the usual partitioning of the
The Hamiltonian

\[ H = F + V, \quad (17) \]

where \( F \) is the zeroth-order Fock Hamiltonian and \( V \) is the remaining part of the usual two-body operator that is taken as first-order, subtracting out the one-body potential contained within the Fock operator. It is customary to treat the Fock \( f_{ia} \) elements between occupied and virtual spaces as first order and other Fock blocks (between occupied-occupied and virtual-virtual spaces) as zeroth order \([30]\). One can then break up the amplitudes into their respective orders, where

\[ T_n = T_n^{(1)} + T_n^{(2)} + T_n^{(3)} + \ldots, \quad (18) \]

and determine their respective order by when these excitations make their first non-zero contribution to the wavefunction. Corrections to double excitations first show up in the first-order wavefunction in MBPT, therefore the first non-zero component to \( T_2 \) occurs at first-order \( T_2^{(1)} \). Corrections to single excitations show up in second-order for Hartree-Fock references and at first-order for non Hartree-Fock references. In the second-order MBPT wavefunction we have contributions from single, double, triple, and quadruple excitations. Hence the triples and quadruples all are taken as second-order.

We can begin by writing out the coupled-cluster equations in the context of coupled-cluster doubles (CCD), where the \( T_1 \) elements are set to zero. For the coupled-cluster \( T_2 \) amplitudes, we expand the orders of the amplitudes and collect terms having the same order, giving
\[
0 = (\Psi_{ij}^{ab} | (F_N T_2 + V_N T_2 + \frac{1}{2} V_N (T_2)^2))_c | \Psi_0 \rangle \\
= (\Psi_{ij}^{ab} | (F_N (T_2^{(1)} + T_2^{(2)} + T_2^{(3)} + ...) + V_N (T_2^{(1)} + T_2^{(1)} + T_2^{(3)} + ...)) \\
+ \frac{1}{2} V_N (T_2^{(1)} + T_2^{(2)} + T_2^{(3)} + ...)^2)_c | \Psi_0 \rangle \\
= (\Psi_{ij}^{ab} | (F_N T_2^{(1)} + (F_N T_2^{(2)} + V_N T_2^{(1)} + (F_N T_2^{(3)} + V_N T_2^{(2)} \\
+ \frac{1}{2} V_N (T_2^{(1)})^2 + ...)_c | \Psi_0 \rangle \\
= (\Psi_{ij}^{ab} | \bar{H}^{(1)} + \bar{H}^{(2)} + \bar{H}^{(3)} + ... | \Psi_0 \rangle ,
\]
where we have partitioned \( \bar{H} \) into its various orders,

\[
\bar{H} = \bar{H}^{(1)} + \bar{H}^{(2)} + \bar{H}^{(3)} + ... 
\]

Of course the elements in the CCSD equation above don’t represent the full \( \bar{H} \) at a given order, but only the terms that can be fully connected with the doubly excited bra wavefunctions. What’s more, by writing it out this way we have that the energies

\[
E^{(n)} = \langle \Psi_0 | \bar{H}^{(n)} | \Psi_0 \rangle 
\]

are just the energies of MBPT(\( n \)). So, starting from a Hartree-Fock reference we have our MBPT(2) energy is just

\[
E^{(2)} = \langle \Psi_0 | \left( V_N T_2^{(1)} \right)_c | \Psi_0 \rangle . 
\]

where from solving the \( T_2 \) amplitude equation above for \( T_2^{(1)} \) we have

\[
0 = \langle \Psi_{ij}^{ab} | \left( F_N T_2^{(1)} \right)_c | \Psi_0 \rangle . 
\]

Working out the diagrams, we see that the equation for \( T_2^{(1)} \) above gives our MP2 amplitudes, and placing these into the second order energy equation above we get the MP2 energy.

One of the primary connections between perturbation theory and the coupled-cluster equations is that the coupled-cluster equations sum up over all infinite orders of
perturbation theory. Looking at the similarity transformed Hamiltonian in the manner above, we can generate all orders of perturbation theory. Differences arise between coupled-cluster and MBPT at a given order based on the truncation of the excitation operator. By solving this latter set of CCSD equations, we are in effect taking into account all infinite perturbation diagrams that are composed of double excitation operators or products of those double excitation operators. In this way we are neglecting perturbations involving the use of triple excitations, quadruple excitations that cannot be written as a product of double excitations, and all higher-order excitations.

### 3.2.2 Perturbative CCSD

As an aside, we will look at the contribution of triples to the MBPT energy. This contribution is particularly important as it will find its way into the CCSD(T) expression. As stated before, the triples are of course second order, but they don’t make their way into the second- and third-order energy expressions $E^{(2)}$ and $E^{(3)}$. Their first appearance is in the fourth order energy $E^{(4)}$. Likewise, the second order quadruples appear in $E^{(4)}$, but the quadruples here are actually products of double excitations and so hence can be described in terms of CCSD. Terms involving the $T_3$ operator additionally shows up in the second-order $\hat{H}^{(2)}$ through the term $(F_N T_3^{(2)})_c$. Ignoring $T_3$ means that CCSD contains the majority of diagrams up to MBPT(4), but is only correct through MBPT(3). This of course makes sense when noticing that the MBPT(2) and MBPT(3) energies only depend on double excitations while single, double, and triple excitations don’t appear until the MBPT(4) correction. Ignoring that preexisting knowledge, we can find that out by perturbative analysis of the coupled-cluster equations. Looking at an expression for determining the $T_3$ amplitudes [31], we can isolate the second-order terms from simple inspection:

$$0 = \langle \psi_{ijk}^{abc} | \hat{H}^{(2)} | \psi_0 \rangle$$

$$0 = \langle \psi_{ijk}^{abc} | (V_N T_2^{(1)} + F_NT_3^{(2)})_c | \psi_0 \rangle$$

(24)

Of course the $T_3$ term doesn’t make its way into the CCSD energy expression since the
$T_3$ can not fully contract with the Hamiltonian and instead affects the energy indirectly through its contribution to $T_2$ and $T_1$. While the $T_3^{(2)}$ technically is contained within the second-order $\tilde{H}^{(2)}$ with its second term above, we have that it won’t be able to enter into the $T_2$ equations until $\tilde{H}^{(3)}$

\[
0 = \langle \Psi_{ij}^{ab} | \tilde{H}^{(3)} | \Psi_0 \rangle \\
0 = \langle \Psi_{ij}^{ab} | \left( F_N T_2^{(3)} + V_N T_1^{(2)} + V_N T_2^{(2)} + \frac{1}{2} V_N (T_2^{(1)})^2 + V_N T_3^{(2)} \right)_c | \Psi_0 \rangle ,
\]

where it enters through the term $(V_N T_3^{(2)})_c$. Of course if we assumed a non-Hartree-Fock determinant one can generate many additional terms that can be found elsewhere \cite{30}.

Again, these terms can be found by first generating all connected diagrams consisting of the doubly excited bra and connected similarity transformed Hamiltonian. Then, assigning the correct order to the Hamiltonian matrix elements, the remaining orders of the $T$ amplitudes can be found by ensuring that the orders of the connected elements add up to a desired perturbation order. Here we have that $F_N$ and $V_N$ are zeroth and first order while the desired perturbation order is three, and thus we can assign the orders of the various $T$ amplitudes like in the equation above. Additionally, we can see how $T_3$ contributes to the $T_1$ term and see the lowest order at which $T_3$ contributes is $\tilde{H}^{(3)}$

\[
0 = \langle \Psi_{ij}^{a} | \tilde{H}^{(3)} | \Psi_0 \rangle \\
0 = \langle \Psi_{ij}^{a} | \left( V_N T_2^{(2)} + V_N T_1^{(2)} + V_N T_2^{(2)} + V_N T_3^{(2)} \right)_c | \Psi_0 \rangle .
\]

Looking at equation 12 however, we see that a third order $T_1^{(3)}$ could only contribute to a fourth-order energy if $f_i^a$ were not zero, as in the case of an arbitrary non Hartree-Fock wavefunction. In this case, however, many of the previous equations would need to be modified as we assumed $T_1$ had its first non-zero elements at second-order.
3.3 Equation of Motion Theory

By performing the similarity transformation on the Hamiltonian and truncating the excitation operator, we have arrived at an effective Hamiltonian \( \hat{H} = e^{-T} H e^{T} \). Just like in CI when you find an approximate wavefunction given an exact Hamiltonian, we can now replace this exact Hamiltonian by an approximate Hamiltonian \( \hat{H} \). Because this Hamiltonian has folded in the effects of higher excitations, we should (hopefully) be able to achieve higher accuracy compared to CI for a given approximate wavefunction.

Thus by viewing the Hamiltonian in this way, we will be able to calculate excited states for this equation-of-motion (EOM) coupled-cluster theory.

Formally, we will be solving the eigenvalue equation

\[
e^{-T} H e^{T} R(m) |\Psi_0\rangle = E R |\Psi_0\rangle \tag{27}
\]

given some excitation operator for state \( m \)

\[
R(m) = R_0(m) + R_1(m) + R_2(m) + ...
\]

By setting \( T = 0 \), we see that we recover just the CI equations. It should be noted that for practical purposes, setting \( T = 0 \) in actual EOM code implementations will not give CI energies since most codes will automatically assume the coupled-cluster equations hold. This assumption of course helps in reducing the cost of the diagonalization in that off diagonal couplings of excited bra states with the reference determinant are assumed to be zero. In addition to this, we can recover full CI for excited states in the limit of infinite excitation operators \( R \), regardless of the order of expansion of \( T \) operators. Since excitation operators commute, one can see the above EOM equation is just using the Hamiltonian \( e^{-T} H R e^{T} \), and since diagonalizing \( HR \) should give the exact full CI spectrum of eigenvalues then these eigenvalues should be preserved under the similarity transform.

The three commonly used are excitation energy (EE) EOM-CCSD, ionization potential (IP) EOM-CCSD, and electron affinity (EA) EOM-CCSD with operators
\[ R_{EE} = \sum_{ia} r_i^a a_i^\dagger a_i + \frac{1}{2} \sum_{ijab} a_i^a b_j^a a_j a_i + ... \] (29)

\[ R_{IP} = \sum_i r_i a_i + \frac{1}{2} \sum_{ij} r_{ij} a_i^\dagger a_j a_i + ... \] (30)

\[ R_{EA} = \sum_a r_i a_i + \frac{1}{2} \sum_{j} r_{j} a_i^\dagger a_j + ... \] (31)

Starting with a similarity transformed \( \hat{H} \) from CCSD, EE EOM-CCSD consists of all excitations up to doubly excited excitations (the \( ph \) and \( 2p2h \) spaces), EOMIP-CCSD consists of all single ionizations along with ionizations and excitations (the \( h \) and \( p2h \) spaces), while EA EOM-CCSD consists of single electron additions along with electron additions and excitations (the \( p \) and \( 2ph \) spaces). In our work we mainly focused on the IP EOM-CCSD and EA EOM-CCSD for their of ionization potentials and electron affinities, and resulting band structure, in solids.

One can define the matrix eigenvalue problem in a similar way to the CI matrix problem,

\[
H = \begin{bmatrix}
E_{CCSD} & \langle \Psi_0 | \hat{H} | S \rangle & \langle \Psi_0 | \hat{H} | D \rangle & \cdots \\
0 & \langle S | \hat{H} | S \rangle & \langle S | \hat{H} | D \rangle & \cdots \\
0 & \langle D | \hat{H} | S \rangle & \langle D | \hat{H} | D \rangle & \cdots \\
\vdots & \vdots & \vdots & \ddots \\
\end{bmatrix}, \quad (32)
\]

where the off-diagonal elements in the first column have been set to zero from satisfying the coupled-cluster equations. Unlike the standard CI matrix equation, the matrix here is not Hermitian and thus both left and right eigenvectors are obtained. The left eigenvectors are defined in an analogous way to the right eigenvectors,

\[ \langle \Psi_0 | L(m) = L_0(m) + L_1(m) + L_2(m) + ... , \] (33)

but rather includes dexcitation operators.
The eigenvalues for the solving the left and right eigenvector equation should be the same, however the left and right eigenvectors will be different. These left and right eigenvectors form a biorthogonal system,

\[ L_n = \frac{1}{(n!)^2} \sum l_{ijk \ldots}^{abc \ldots} a_i^+ a_j^+ a_k^+ \cdots a_c a_b a_a. \]  

(34)

where the normalization constant \( C \) is chosen to be unity to satisfy the ground state energy equation [32]. This is done by first by normalizing the right eigenvector and subsequently normalizing the left eigenvector such that the above equality holds. It should be noted that besides the normalization constant, the orthogonality property should hold between different states \( m \) and \( n \) if the procedure is properly implemented. Similar to the case of CI, the size of the eigenvalue equation usually calls for an iterative eigensolver due to issues of computational time and memory. For this, typically a non-Hermitian Davidson procedure and its variants are used for finding the eigenvalues [33, 34].

### 3.4 EOM-CCSD Equations

Here we will show the ionization potential (IP) and electron affinity (EA) EOM-CCSD equations for both the right eigenvectors and left eigenvectors done using a spatial orbital basis. While the left and right eigenvector solutions should necessarily give the same exited state eigenvalues, the eigenvectors should be different due to the non-Hermitian operator \( \tilde{H} \). The left eigenvectors are needed for property calculations where expectation values are needed, and more importantly for perturbation theories where we will try to obtain perturbations to the EOM-CCSD eigenvalues. It is beneficial to show these various eigenvector equations in a single document for reasons of clarity; many of the documented equations use various differing basis sets for the left and right EOM eigenvectors. This leads to more differences in the structure of the IP and EA.
equations, when in reality they are quite similar. This should come as no surprise as one theory simply has a hole operator in place, for the IP, of the other’s particle operator, for the EA, and vice versa.

### 3.4.1 IP EOM-CCSD Equations

The right and left eigenvector equations in closed-shell spatial orbital form for the ionization potential are given in equations 36 and 37, respectively.

\[
(\hat{H}R)_i = -\sum_k U_{ki} r_k + \sum_{ld} U_{ld} (2r^d_{il} - r^d_{li}) - \sum_{kld} (2W_{kld} - W_{lkid}) r^d_{kl} \\
(\hat{H}R)^b_{ij} = \sum_d U_{bd} r^d_{ij} - \sum_l U_{li} r^b_{lj} + \sum_{il} U_{lj} r^b_{il} - \sum_{ld} W_{lbj} r^d_{ij} - \sum_{ld} W_{lbj} r^d_{ij} \\
+ \sum_{kl} W_{klij} r^b_{kl} - \sum_{kd} W_{klij} r^b_{kl} - \sum_c \left[ \sum_{kld} (2W_{lkdc} - W_{lkdc}) r^d_{kl} \right] t^b_{ij}
\]

One can immediately see that the summation terms 5-8 in equation 36 have a cost of $N^5$, specifically $n_v^2 n_o^3$ with the exception of summation 7 which is $n_v n_o^4$. Moving on to the left eigenvector equation, we have
\[(LH)_i = - \sum_k U_{k,l}^i - \sum_{k,b} W_{k,bj}^l l_{ij}^b \]

\[ (LH)^d_{kl} = 2U_{ld,k}^i - \sum_i (2W_{klid} - W_{lkid}) l_i^d \]

\[
\begin{align*}
&+ \sum_b U_{bd,k}^l - \sum_i U_{kd,li}^d - \sum_j U_{lj}^d k_j^d \\
&+ \sum_{bj} (2W_{lbjd} - W_{lbjd}) l_{kj}^b - \sum_{bj} W_{kbjd} l_{ij}^b \\
&+ \sum_{ij} W_{klij} l_{ij}^d - \sum_{ba} W_{kbdl} l_{id}^b \\
&- \sum_c (2W_{lkde} - W_{klic}) \sum_{ijb} t_{bij} l_{ij}^b
\end{align*}
\]

(37)

Transforming the right EOM equations to obtain the equations for the left eigenvector is relatively straightforward. For each of the diagrams in the right EOM equations, one obtains a diagram with open lines at the top, to connect to the left eigenvector basis, and the right amplitudes at the bottom of the diagram. The general form of these diagrams can be seen in the article by Nooijen and Bartlett [35]. To form the left eigenvector equation, one instead uses the open diagram at the top to be the left amplitudes and opens the lines at the bottom of the diagram. Now the top of the diagram will correspond to the left amplitudes and the bottom of the diagram will be able to connect with the right eigenvector basis.

Or more simply from a matrix multiplication point of view, our right hand equation is just of the general form \(A_{ij} r_j \rightarrow (AR)_i\). The left hand multiplication is just \(l_j^T A_{ij} \rightarrow (LA)_j^T\) or simply \(A_{ji} l_i \rightarrow (LA)_j\). Given the term \((AR)_i\) above, we can replace the right eigenvectors by left eigenvectors, and quite simply we switch the contracted indices with the free indices, and our final result should contribute to \((LA)_j\). To give an example, we first look at \((HR)_i\) and take the first term in the summation from the second line,

\[- \sum_{kld} 2W_{kld} r_{kd}^d \rightarrow (HR)_i.\]

From our notation above, see our free index is \(i = \{i\}\) and our contraction indices are
Thus our final expression will involve a contribution to $(L\tilde{H})_{kl}^d$, from the contraction indices, and involve summation over $l$, found from the free indices. This term will involve contraction with the index $l_i$, since that was the index for the result $(\tilde{H}R)_i$. Putting this together we have

$$-\sum_i 2W_{klid}l_i \rightarrow (L\tilde{H})_{kl}^d.$$  

One can see this term shows up as the first term in the summation for the $(L\tilde{H})_{kl}^d$ terms above. Thus given the right eigenvector equations one has an easy way to obtain the left-hand eigenvector equation without having to redo (or even write) any diagrams.

### 3.4.2 EA EOM-CCSD Equations

The right eigenvector equation and left eigenvector equations in closed-shell spatial orbital form for the electron affinity are given in equations 38 and 39, respectively.

\[
(\tilde{H}R)^a = \sum_c U_{ac}r_c^a + \sum_{ld} U_{id}(2r_{l}^{ad} - r_{l}^{da}) \\
+ \sum_{lcd} (2W_{alcd} - W_{aldc})r_{l}^{cd} \\
(\tilde{H}R)^{ab}_{lj} = -\sum_l U_{lj}r_{l}^{ab} + \sum_c U_{acr_{j}^{ab}} + \sum_d U_{bd}r_{l}^{ad} \\
+ \sum_c W_{abcj}r_{c}^{ab} + \sum_{ld} (2W_{lbdj} - W_{bdlj})r_{l}^{ad} - \sum_{lc} W_{abcj}r_{l}^{cb} \\
+ \sum_{cd} W_{abcd}r_{j}^{cd} - \sum_{lc} W_{blij}r_{l}^{ca} \\
- \sum_k \left[ \sum_{lcd} (2W_{kldc} - W_{klcd})r_{k}^{cd} \right] l_{kj}
\]

(38)

One can immediately see that the summation terms 5-8 in equation 38 have a cost of $N^5$, specifically $n^3_vn^2_o$ with the exception of summation 7 which is $n^4_en_o$. 

30
\[(LH)^c = \sum_a U_{ac} l^a + \sum_{jab} W_{abcj} l^a j\]

\[(LH)_{l}^{cd} = 2U_{ld} l^c - U_{lc} l^d + \sum_a (2W_{abcd} - W_{alcd}) l^a\]

\[+ \sum_a U_{ac} l^a + \sum_b U_{bd} l^b + \sum_j U_{ij} l^d j\]

\[+ \sum_{bj} (2W_{lbjd} - W_{lbjd}) l^b j - \sum_{aj} W_{laje} l^d j\]

\[+ \sum_{ab} W_{abcd} l^a - \sum_{bj} W_{lbcj} l^b j\]

\[\sum_j (2W_{kjfe} - W_{kjfe}) \sum_{ibc} t_{ij} e_{ijl} \]

\[3.5 \text{ Perturbative Triples: EOM-CCSD*}\]

These left and right eigenvectors can then be used in perturbation theory. While numerous perturbation theories exist for the excited state [36–39], we focused on the EOM-CCSD* approach by Stanton and coworkers for a perturbative treatment of triples due to its straightforward theory. The EOM-CCSD* method was proposed by Stanton and coworkers [40–42] for a perturbative treatment of triples for EOM-CCSD calculations. The correction is a non-iterative $O(N^6)$ correction, adding a factor of $N$ to the cost of the EOM-CCSD calculation but having the same order as the underlying CCSD ground state calculation. Unfortunately, the correction and its variations does not give the same kcal/mol accuracy as the perturbative triples treatment for ground state calculations, CCSD(T).

The equations are given by [40]:

\[
D_{ijk}^{ab} = P(ijk)^{ab} (ab \mid ij) - P(ijk) \sum_e l^e_{ij} \langle ek \mid ab \rangle - P(ab) P(ijk) \sum_m l^a_{mk} \langle ij \mid mb \rangle
\]
\[ D_{ijk}^{ab} r_{ij}^{ab} = -P(ijk) \sum_e r_{ij}^{e} \langle ab | ek \rangle - P(ab)P(ijk) \sum_m r_{mk}^a \langle mb | ij \rangle - P(ab)P(ijk) \sum_{me} r_{mi}^{ae} \langle mb | ke \rangle + P(kji) \sum_{mn} r_{mn}^{ab} \langle mn | kj \rangle - P(ijk) \sum_m r_{mk}^e [P(ij) \sum_{en} r_{nj}^{ab} \langle mn | ei \rangle ] + P(ijk)P(ab) \sum_{mef} r_{mk}^{ef} t_{ij}^{lb} \langle ma | ef \rangle + \frac{1}{2} P(ijk)P(ab) \sum_{mnc} r_{mn}^{a} r_{ij}^{eb} \langle mn | ek \rangle \] (41)

An additional four terms were excluded from the original formulation as well that will be added here:

\[ D_{ijk}^{ab} r_{ij}^{ab} \approx -P(ijk) \sum_m r_{mk}^{b} [P(ij) \sum_{en} r_{nj}^{a} \langle mn | ei \rangle ] - P(ijk)P(ab) \sum_{mef} t_{mk}^{eb} t_{ij}^{fi} \langle ma | ef \rangle - \frac{1}{2} P(ijk)P(ab) \sum_{mnc} r_{mn}^{ab} r_{ij}^{ce} \langle mn | ek \rangle \] (42)

In the redefined formulation the final three terms in the first \( r_{ij}^{ab} \) equation and the three terms in the second \( r_{ij}^{ab} \) above equation were excluded with no noticeable drop in accuracy. One reason for this is that these terms correspond to a higher order of perturbation theory than those of the other terms. Besides that, it was also found that these two terms had a cancelling effect when added together and gave a modest improvement to the accuracy [40].
4 Periodic Systems

This section is here to develop some of the basics of doing electronic structure in solids. Unless one takes a solid state physics course, and even for those that have, the average quantum chemist generally has little exposure to the solid state, in particular the derivation of the matrix elements that go into creating a program for mean-field or correlated electronic structure methods. While there are various online tutorials and websites to help people in this regard, there has not been many formal introductions to the subject including the one I found most helpful by Evarestov [43]. I take a more practical approach here and derive the Hamiltonian matrix elements for Gaussians expanded in terms of planewaves. But by understanding the concepts here, namely the Fourier transforms and their properties, one should be able to derive any new matrix elements that are not presented here for either ab-initio solid state systems or even model systems such as the jellium.

We will first explore the concept of unit cells and supercells and what it means to be a periodic system. Next we will explore some of the properties of the basis functions used in periodic systems and look into the periodic symmetry as encoded in k-points. Then we will talk a bit about the Fourier transform and its identities that will be useful later in deriving the matrix elements. I will discuss momentum conservation and how it
applies to the periodic Hamiltonian to help reduce the number of matrix elements that
need to be calculated and stored. Finally we will derive the one and two-body matrix
elements and will hit on topics such as the Fourier Poisson equation and aliasing for
electronic integrals.

4.1 The Basics of Periodic Systems

4.1.1 Unit Cells and Supercells

Without the use of symmetry, the solution of the Schrödinger equation in solids would
be a near impossible task, or at the very least require much more sophisticated ma-
chinery when solving even the simplest of mean-field Hamiltonians. The development
of electronic structure methods in quantum chemistry has relied on the translational
symmetry of the Hamiltonian and hence wavefunction. The translational symmetry
comes about from the assumption of periodic boundary conditions; while a true crystal
has a large finite size, the periodic boundary conditions (PBC) assumes an truly infinite
crystal composed of a repeating unit cell along some periodic axis or axes. With this
assumption comes simplifications. While in the large finite crystal case, we could have
that two crystals could feel two different potentials depending on its orientation within
the crystal structure. Unit cells near the surface might feel very different potentials
from those within the bulk of the material. With PBC we can say the potential felt
in a unit cell from the entirety of the crystal, i.e. from all other infinitely repeating
unit cells, is the same as the potential felt in another unit cell from the entirety of the
crystal. In some sense, the unit cell we choose as an origin is insignificant. We can
solve for the energy of one unit cell and it should be the same regardless of any other
unit cell chosen. Adding in the energy of any additional cells only serves to multiply
that energy by an integer factor.

We will now develop this slightly more formally, yet not completely rigorous (and
hopefully not too dry!). Given unit cell translational vectors $\mathbf{a}_1, \mathbf{a}_2, \mathbf{a}_3$ and integers
$m_1, m_2, m_3$ we can construct an arbitrary translational vector $\mathbf{R}$. 

34
\[ \mathbf{R} = m_1 \mathbf{a}_1 + m_2 \mathbf{a}_2 + m_3 \mathbf{a}_3. \] (43)

With these translational vectors, we can translate our unit cell by the set of all \( \mathbf{R} \) and generate the entire crystal. This set of all translations \( \{ \mathbf{R} \} \) is called the Bravais lattice. The crux of our analysis is that the potential is periodic, which we will assume to hold true, and that given any point \( \mathbf{r} \) we have

\[ V(\mathbf{r} + \mathbf{R}) = V(\mathbf{r}). \] (44)

Most potentials in simple molecular systems would of course not satisfy this property; but more curiously, the presence of an electric field would also not satisfy this property, regardless of whether in a molecular of periodic system. We will generally restrict \( \mathbf{r} \) to be within the unit cell unless otherwise explicitly stated. We define a point to be within our unit cell if it can be written as \( \sum_{i=1}^{3} b_i \mathbf{a}_i \) for \( b_i \in [0, 1) \).

Now comes the question of how we choose the unit cell in the first place. There are of course many different choices: 1) the \textit{primitive cell} that is the smallest possible unit cell, 2) the \\textit{Wigner-Seitz cell} is a primitive cell that contains all points in space that are closer to a given Bravais lattice point that any other, and lastly, 3) any other unit cell that can create the entire crystal lattice without any overlapping between neighboring cells through translation. The only choice here that is unique is the Wigner-Seitz cell and it has the special property that all points within the cell are so-called minimum images to the Bravais lattice point on which it is centered, by construction.

It should be somewhat obvious that one could define a unit cell to be a collection of smaller unit cells. We can do this in the following way: choose positive integers \( N_i \) and define new translation vectors

\[ \mathbf{R} = N_1 \mathbf{a}_1 + N_2 \mathbf{a}_2 + N_3 \mathbf{a}_3. \] (45)

We call this unit cell the \textit{supercell}. This will be the basic method for reaching the thermodynamic limit, where the bulk of the infinite system will be modelled. The
concept of a k-point will be used to model the supercell efficiently, using the symmetries inherent in the cell structure. We define our wavefunction to be periodic over the supercell and for this reason one can see why a calculation done with a single unit cell will not be sufficient to accurately describe the properties of the system under study. The reasons for this can be seen in the off-diagonal of the density matrix $\rho(r, r')$, where due to the PBC we have an artificial periodicity introduced into the density

$$\rho(r + R, r' + R) = \rho(r, r'). \tag{46}$$

This off-diagonal element should instead decay exponentially with respect to distance for semiconductors [44]. Even if we were to simply cut off the interaction at a certain distance $R$, we would still not be able to capture the full extent of this exponential tail correctly if $R$ is too small. By increasing the size of the supercell we can systematically approach the thermodynamic limit, typically reaching the thermodynamic properties we desire with supercell sizes of roughly $4 \times 4 \times 4$ (composed of a repeating unit of four unit cells in each periodic direction) for semiconductors and $7 \times 7 \times 7$ for metals. This larger size for metals is tied with the fact that off-diagonal elements decay only polynomially in metallic systems.

### 4.1.2 Planewaves and k-points

The symmetry in periodic calculations is captured in Bloch’s Theorem, which states that the wavefunction for any periodic potential can be written as

$$\psi_k(r) = u_k(r)e^{ik \cdot r} \tag{47}$$

where $u_k(r)$ is a function periodic over the unit cell, i.e.

$$u_k(r + R) = u_k(r). \tag{48}$$

One obvious choice for a basis function that satisfies this periodicity condition is the planewave basis, those of the form $e^{iG \cdot r}$, where the $G$ are written as linear integer
combinations of the reciprocal lattice vectors

\[ \mathbf{G} = \{m_1 \mathbf{G}_1 + m_2 \mathbf{G}_2 + m_3 \mathbf{G}_3\}, \quad (49) \]

where the lattice vectors satisfy the biorthogonality condition

\[ \mathbf{G}_i \cdot \mathbf{a}_j = 2\pi \delta_{ij}. \quad (50) \]

Another way of rewriting the equation above is to say

\[ \mathbf{G} \cdot \mathbf{R} \equiv 0 \pmod{2\pi}. \quad (51) \]

Because of the biorthogonality, the planewaves can be shown to be periodic over the unit cell, and so we can let the cell-periodic part be represented by planewaves

\[ u_k(\mathbf{r}) = e^{i\mathbf{G} \cdot \mathbf{r}}. \quad (52) \]

The reciprocal vectors satisfying 50 are found by solving the simple linear equation \( \mathbf{G} \mathbf{A}^T = 2\pi \mathbf{1} \), where the rows of \( \mathbf{G} \) and \( \mathbf{A} \) are the \( n \)-dimensional reciprocal and translational vectors, respectively, of the given system and \( \mathbf{1} \) is the identity matrix. Since one cannot include the full (infinite) set of planewave in a calculation, one needs a way to include a finite number of planewaves in a calculation and approach the complete basis set limit. To do this, one includes only a subset of planewaves up to some maximum frequency. One common way to include all elements with kinetic energy less than some given energy cut-off \( E_{\text{cut}} \),

\[ \frac{1}{2} |\mathbf{k} + \mathbf{G}|^2 < E_{\text{cut}}. \quad (53) \]

Another way, which is done in the PySCF [10] is to include the set of all planewaves with a cut-off in each periodic direction,

\[ \left\{ \sum_{i=1}^{3} m_i \mathbf{G}_i \text{ where } |m_i| \leq \tilde{m}_i \right\} \quad (54) \]
given $\tilde{m}_1, \tilde{m}_2, \tilde{m}_3$. These maximum cut-offs are usually defined so that the maximum momenta in any one direction is less than a given kinetic energy cut-off.

Moving beyond planewaves, another common basis set used in electronic structure theory is the Gaussian basis set. We can see a Gaussian function $\tilde{\phi}_{pk}(r)$ placed in a single unit cell would not immediately satisfy the condition in 48 since it is not translationally invariant. We can instead create a translationally invariant Gaussian-like function using a sum of periodic images plus a phase-factor

$$u_{pk}(r) = \sum_{R} \tilde{\phi}_{p}(r - R)e^{-ik(r-R)}, \quad (55)$$

where the summation is formally over the full set of periodic images. We note the summation is convergent since the Gaussians decay exponentially with distance. Since the summation is over the full set of periodic images, we have that

$$u_{pk}(r + R') = \sum_{R} \tilde{\phi}_{p}(r + R' - R)e^{-ik(r+R'-R)}$$

$$= \sum_{R''} \tilde{\phi}_{p}(r + R'')e^{-ik(r+R'')}$$

$$= u_{pk}(r) \quad (56)$$

where we have defined $R'' = R - R'$ and used the fact that this new set $\{R''\}$ is the same as the set $\{R\}$ (since translating a periodic lattice by a lattice translation vector still recovers the full lattice). We see that this new function is periodic over the unit cell. The full Bloch wavefunction for this orbital, after multiplication by $e^{ikr}$, is

$$\phi_{pk}(r) = \sum_{R} \tilde{\phi}_{p}(r - R)e^{ikR}. \quad (57)$$

This formally is the basis used in PySCF. However, the Hamiltonian matrix elements will generally involve expanding these over a set of planewaves as the matrix elements are more easily computed when done in reciprocal space.

We can easily another consequence of Bloch’s Theorem:
\[ \Psi_k(r + R) = \Psi_k(r)e^{ikR}. \]  

(58)

This corollary is perhaps slightly easier to see (in a non-rigorous way) \(^1\). We know that the density should be periodic, otherwise a mean-field theory such as Hartree-Fock would give different potentials depending on the choice of unit cell, which would violate the assumption that the potential was periodic. Because of this, the density should remain invariant if translated by a translation vector \( R \),

\[ |\Psi_k(r + R)|^2 = |\Psi_k(r)|^2. \]  

(59)

We can try to relate the two wavefunctions by some constant depending on the translation vector, i.e.

\[ \Psi_k(r + R) = C_k(R)\Psi_k(r). \]  

(60)

Since the density norms above are equal, we see that this factor must have length one,

\[ |C_k(R)|^2 = 1. \]  

(61)

We can then write this factor relating the wavefunction and its translation as an exponential \( C_k(R) = e^{ikR} \), giving

\[ \Psi_k(r + R) = e^{ikR}\Psi_k(r). \]  

(62)

But we aren’t done yet; there’s still the task of finding exactly what these \( k \) correspond to. By forcing the wavefunction repeats itself after \( N_i \) unit cells, where \( i \in \{1, 2, 3\} \), we find that only a certain set of \( k \) can be allowed for this condition to hold true. This other condition of periodicity, using \( R = \sum_{i=1}^{3} N_i a_i \), is

\(^1\)More formally, we would start with the fact that the Hamiltonian \( \hat{H}(r) \) commutes with a translation operator \( \hat{T}_R \) and so the wavefunction must be an eigenfunction of the translation operator and be invariant to the translation operator up to a multiplicative constant.
\[
\Psi_k \left( r + \sum_{i=1}^{3} N_i a_i \right) = e^{i k \cdot \left( \sum_{i=1}^{3} N_i a_i \right)} \Psi_k(r) = \Psi_k(r).
\]

(63)

For this to hold true, we must have \( e^{i k \cdot \left( \sum_{i=1}^{3} N_i a_i \right)} = 1 \), which defines our set of k-points. In total, this will define a set of \( N_1 N_2 N_3 \) k-points given by

\[
k = \left\{ \frac{n_1}{N_1} G_1 + \frac{n_2}{N_2} G_2 + \frac{n_3}{N_3} G_3 \right\}
\]

(64)

where \( 0 \leq n_i < N_i \). Just as the largest reciprocal vectors give us a sense of the smallest real space resolution, the smallest k-points in each direction gives us a sense of the largest real space resolution. Smaller k-points will mean larger \( N_i \), and hence the norms of the smallest basis vectors in k-space is innately tied to our assumption about the (enforced) periodicity of our wavefunction.

The supercell containing \( N_i \) cells will give the same energy per unit cell if done at the same k-points, and commensurate basis set, as those defined above. Mathematically, the two differ merely by a unitary transformation of the Hamiltonian. In the case of k-points, this produces a block diagonal structure, where now each of these smaller Hamiltonian blocks labelled by a symmetry \( k \) can be diagonalized independently.

For the case of the supercell with the Hartree-Fock Hamiltonian, diagonalization of the Fock matrix would correspond to diagonalizing an \((N_{cells} n_{mo}) \times (N_{cells} n_{mo})\) matrix, where \( N_{cells} \) is equal to \( N_1 N_2 N_3 \) and \( n_{mo} \) is the number of molecular orbitals per unit cell. For the case of k-points, we would instead diagonalize \( N_{kpt} \) (= \( N_{cell} \)) different Fock matrices each labelled by a k-point, where each Fock matrix is now of size \((n_{mo}) \times (n_{mo})\). Since diagonalization of an \( N \times N \) matrix is \( \mathcal{O}(N^3) \), we have that the cost for the supercell is \( \mathcal{O}(N_{cells}^3 n_{mo}^3) \) while for the k-point version it is merely \( \mathcal{O}(N_{cells} n_{mo}^3) \), giving a savings of \( N_{kpt}^2 \).
4.1.3 Fourier Transforms and Useful Identities

Before getting into our discussion of the matrix elements, it helps just to look at some elementary properties of Fourier integrals. Given a function \( f(\mathbf{r}) \), we can find the components in reciprocal space with

\[
f(\mathbf{G}) = \int_V d\mathbf{r} f(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}} = \frac{V}{N_G} \sum_{\mathbf{r}} f(\mathbf{r}) e^{-i\mathbf{G}\cdot\mathbf{r}} = \frac{V}{N_G} \text{FFT}[f(\mathbf{r})](\mathbf{G})
\]

(65)

\[
f(\mathbf{r}) = \frac{1}{N_G} \sum_{\mathbf{G}} f(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}} = \text{IFFT}[f(\mathbf{G})](\mathbf{r})
\]

(66)

where \( N_G \) is the number of reciprocal vectors, \( V \) is the volume of the unit cell, \( \text{FFT} \) is the Fourier transform, and \( \text{IFFT} \) is the inverse Fourier transform. The factor in front of the Fourier transform in equation 65 is the so-called “sampling interval” and takes into account the fact we integrating over a finite region of space, the unit cell. There are various useful identities when working with discrete Fourier transforms, the first being

\[
\sum_{\mathbf{G}} e^{-i\mathbf{G}\cdot(\mathbf{r}'-\mathbf{r})} = N_G \delta_{\mathbf{r},\mathbf{r}'}.
\]

(67)

Next there is a less used version (at least for the subsequent derivations) for the forward Fourier transform,

\[
\sum_{\mathbf{r}} e^{i(\mathbf{G} - \mathbf{G}')\cdot\mathbf{r}} = N_G \delta_{\mathbf{G},\mathbf{G}'}.
\]

(68)

The more commonly used version of this is

\[
\int d\mathbf{r} e^{i(\mathbf{G} - \mathbf{G}')\cdot\mathbf{r}} = V \delta_{\mathbf{G},\mathbf{G}'}.
\]

(69)

For the proof of the first two Fourier identities, we can see the summations are merely
over the roots of unity. The roots of unity will add up to zero unless one sums up the zeroth root of unity, resulting in a factor of $N_G$. The final equality is easiest seen in that the planewaves are simply orthogonal over the unit cell. These identities will be used ubiquitously throughout the derivation of the Hamiltonian matrix elements.

### 4.2 Periodic Hamiltonian and Momentum Conservation

Using the result in 55, the cell-periodic part of the bloch wavefunction and the full bloch wavefunction (after multiplication by $e^{i k \cdot r}$) for orbital $\phi_{pk}(r)$ is written as

$$ u_k(r) = \sum_R \tilde{\phi}_p(r-R) e^{-ik \cdot (r-R)} $$

$$ \phi_{pk}(r) = e^{ikR} u_k(r) $$

$$ = \sum_R \tilde{\phi}_p(r-R) e^{ikR} $$

(70)

where $\tilde{\phi}(r)$ are the molecular (non-periodic) Gaussian orbitals. The fourier transform of the cell periodic part is simply

$$ u_k(G) = \int dr \ u_k(r) e^{-i G \cdot r} $$

$$ = \int dr \ e^{-i (k+G) \cdot r} \sum_R \tilde{\phi}_p(r-R) e^{i k R}. $$

(71)

And of course we can rewrite our real space quantities in 70 using an inverse fourier transform of this newly defined quantity

$$ u_k(r) = \frac{1}{N_G} \sum_G u_{pk}(G) e^{i G \cdot r} $$

$$ \phi_{pk}(r) = \frac{1}{N_G} \sum_G u_{pk}(G) e^{i (k+G) \cdot r}. $$

(72)

where $N_G$ is the number of basis vectors and the last equality is simply the fact that $\phi_{pk}(r) = e^{i k \cdot r} u_k(r)$. Now one can construct the different elements of the Hamiltonian.

#### 4.2.1 Overlap Matrix Elements

First we look at the overlap matrix element $S_{p\mu,\bar{p}q \sigma}$. 


\[ S_{\mathbf{k}_p\mathbf{k}_q} = \langle \phi_{\mathbf{k}_p}(\mathbf{r}) | \phi_{\mathbf{k}_q}(\mathbf{r}) \rangle \]
\[ = \frac{1}{N_G^2} \int d\mathbf{r} \sum_{G,G'} u^*_{\mathbf{k}_p}(G) u_{\mathbf{k}_q}(G') e^{i(\mathbf{k}_q - \mathbf{k}_p + G - G') \cdot \mathbf{r}} \]
\[ = \frac{1}{N_G^2} \sum_{G,G'} u^*_{\mathbf{k}_p}(G) u_{\mathbf{k}_q}(G') \int d\mathbf{r} e^{i(\mathbf{k}_q - \mathbf{k}_p + G - G') \cdot \mathbf{r}} \]
\[ = \frac{V}{N_G^2} \sum_{G,G'} u^*_{\mathbf{k}_p}(G) u_{\mathbf{k}_q}(G') \delta_{\mathbf{k}_q, \mathbf{k}_p + G - G'} \]
\[ = \frac{V}{N_G^2} \sum_{G} u^*_{\mathbf{k}_p}(G) u_{\mathbf{k}_q}(G), \] (73)

where in the last line we noted that

\[ u_{\mathbf{k}_q}(G') \delta_{\mathbf{k}_q, \mathbf{k}_p + G - G'} = \left[ \int d\mathbf{r} e^{-i(\mathbf{k}_q + G') \cdot \mathbf{r}} \sum_{\mathbf{R}} \phi'_q(\mathbf{r} - \mathbf{R}) e^{i\mathbf{k}_q \cdot \mathbf{R}} \right] \delta_{\mathbf{k}_q, \mathbf{k}_p + G - G'} \]
\[ = \int d\mathbf{r} e^{-i(\mathbf{k}_p + G - G') \cdot \mathbf{r}} \sum_{\mathbf{R}} \phi'_q(\mathbf{r} - \mathbf{R}) e^{i(\mathbf{k}_p + G - G') \cdot \mathbf{R}} \]
\[ = \int d\mathbf{r} e^{-i(\mathbf{k}_p - G) \cdot \mathbf{r}} \sum_{\mathbf{R}} \phi'_q(\mathbf{r} - \mathbf{R}) e^{i\mathbf{k}_p \cdot \mathbf{R}}, \] (74)

and we used the fact that \((G - G') \cdot \mathbf{R} \equiv 0 \pmod{2\pi} \).

One of the importance consequences of symmetry is that we see our overlap matrix now only depends on one of the \(k\)-points,

\[ S_{\mathbf{k}_p\mathbf{k}_q} = \frac{V}{N_G^2} \sum_{G} u^*_{\mathbf{k}_p}(G) u_{\mathbf{k}_q}(G) \delta_{\mathbf{k}_p, \mathbf{k}_q}, \] (75)

which shows that only orbitals with the same symmetry, i.e. that are at the same \(k\)-point, can have overlap. The same can be shown for one-electron operators, where matrix elements between two determinants can be non-zero only if the two determinants are at the same \(k\)-point, i.e. \(\mathbf{k}_p - \mathbf{k}_q \equiv 0 \pmod{\mathbf{G}}\), which is a consequence of momentum conservation arising from group theory and symmetry. Here the \(\mathbf{k}_p - \mathbf{k}_q \equiv 0 \pmod{\mathbf{G}}\) means that there exists some \(m_1, m_2, m_3\) such that

\[ \mathbf{k}_p - \mathbf{k}_q + \sum_{i=1}^{3} m_i \mathbf{G}_i = 0, \] (76)
or in other words, that the difference between the k-points is equivalent to zero when translated into the first Brillouin zone.

For two-electron operators like the MO integral \( \langle p | k_p | q | r | k_r | s | k_s \rangle \) the analogous condition is

\[
\mathbf{k}_p + \mathbf{k}_q - \mathbf{k}_r - \mathbf{k}_s \equiv 0 \quad (\text{mod } \mathbf{G}).
\]  
(77)

The modulo part is especially important, as the difference of k-points can bring them outside the so-called “first Brillouin zone” and so this needs to be accounted for in practice by finding the integer \((m_1, m_2, m_3)\) such that this modular condition holds. The main consequence is that for one-electron matrix elements, one of the k-points is “free” while the other is fixed, and for two-electron matrix elements, we have that three of the k-points are free while the final one is fixed from momentum conservation. This reduces both the computational complexity as well as the memory requirements for the calculation. Since many of the calculations can have upwards of 64 to even hundreds of k-points, this reduction in cost can be quite substantial.

### 4.2.2 Kinetic Energy Matrix Elements

The kinetic energy matrix element is given by

\[
K_{kp,pq} = -\frac{1}{2} \langle \phi_{kp}(\mathbf{r}) | \nabla^2 | \phi_{pq}(\mathbf{r}) \rangle
= -\frac{1}{2N_G^2} \sum_{G,G'} u^*_{kp}(G) u_{pq}(G') \int d\mathbf{r} e^{-i(k_p+G)\mathbf{r}} \nabla^2 e^{-i(k_p+G')\mathbf{r}}
\]

\[
\begin{align*}
&= \frac{1}{2N_G^2} \sum_{G,G'} u^*_{kp}(G) u_{pq}(G') |k_p + G'|^2 \int d\mathbf{r} e^{-i(G'-G)\mathbf{r}} \\
&= \frac{V}{2N_G^2} \sum_{G,G'} u^*_{kp}(G) u_{pq}(G') |k_p + G'|^2 \delta_{G,G'} \\
&= \frac{V}{2N_G^2} \sum_{G} u^*_{kp}(G) u_{pq}(G) |k_p + G|^2
\end{align*}
\]  
(78)
4.2.3 Electron-Nuclear Matrix Elements

The Poisson’s equation describes the potential, $V(\mathbf{r})$, for a given charge distribution, $\rho(\mathbf{r})$, through the following differential equation written in atomic units

$$\nabla^2 V(\mathbf{r}) = -4\pi \rho(\mathbf{r}). \quad (79)$$

We can expand both the real-space potential and charge in terms of its Fourier components, i.e. for the potential this is $V(\mathbf{r}) = \frac{1}{V} \sum_{\mathbf{G}} V(\mathbf{G}) e^{i\mathbf{G}\cdot\mathbf{r}}$ and an analogous equation holds for the density, and substitute this into the Poisson equation above. Note that for the Fourier transform of the potential, this is not the result of a discrete Fourier transform and so the normalization factor of $1/V$ is used instead of $1/N_G$.

In doing so, we arrive at the Fourier Poisson’s equation, where we describe the potential terms in reciprocal space through the Fourier components of the density,

$$V(\mathbf{G}) = \frac{4\pi \rho(\mathbf{G})}{|\mathbf{G}|^2}. \quad (80)$$

While this was the result of a true Fourier transform (to distinguish it from a discrete Fourier transform), we will however usually use a $\rho(\mathbf{G})$ found from a discrete Fourier transform as there has been much theory and work involved in their fast implementation. For the specific case of the electron-nuclear potential due to an atom $A$, we write the usual Fourier transform for a potential $1/|\mathbf{r}|$ but substitute $(\mathbf{r} - \mathbf{r}_A)$ for $\mathbf{r}$,

$$V_A^{en}(\mathbf{G}) = \int d\mathbf{r} V_A^{en}(\mathbf{r}) e^{-i\mathbf{G}\cdot(\mathbf{r} - \mathbf{r}_A)} = \frac{4\pi Z_A}{|\mathbf{G}|^2}. \quad (81)$$

In the above equation it is understood that $V_A^{en}(\mathbf{r}) = Z_A/|\mathbf{r} - \mathbf{r}_A|$. We do the same for the Fourier component of the potential, letting

$$V_A^{en}(\mathbf{r}) = \frac{1}{V} \sum_{\mathbf{G}} V(\mathbf{G}) e^{i\mathbf{G}\cdot(\mathbf{r} - \mathbf{r}_A)}. \quad (82)$$

In most periodic calculations one will find pseudopotentials used instead of the bare
nuclear potential as the bare nuclear potential needs a high number of planewaves to represent properly. This discussion can be found elsewhere [45] and will not be discussed here as there has been much research and history dedicated to this subject. For now, however, it suffices to say that we used the Goedecker-Teter-Hutter (GTH) pseudopotentials along with corresponding basis sets optimized for these pseudopotentials as found in CP2K [1, 46, 47]. We will now derive the electron-nuclear attraction matrix elements.

The electron-nuclear attraction matrix element, \( V_{en}^{\text{kn}} \), can be written as

\[
V_{en}^{\text{kn}} = -\langle \phi_{kp}(r) | \sum_A \frac{Z_A}{r - r_A} | \phi_{kq}(r) \rangle 
\]

\[
= \frac{1}{VN_G^2} \sum_{G,G'} u_{kp}^{*}(G) u_{kq}^{*}(G') \sum_A \sum_{G''} V_{A}^{en}(G'') \int dr e^{i(kp+G' - kq - G)r} e^{iG''(r - r_A)} 
\]

\[
= \frac{1}{VN_G^2} \sum_{G,G'} u_{kp}^{*}(G) u_{kq}^{*}(G') \sum_A \sum_{G''} V_{A}^{en}(G'') e^{-iG''r_A} \int dr e^{i(G' - G + G'')r} 
\]

\[
= \frac{1}{N_G^2} \sum_{G,G'} u_{kp}^{*}(G) u_{kq}^{*}(G') \sum_A \sum_{G''} V_{A}^{en}(G'') e^{-iG''r_A} \delta_{G''G - G'} 
\]

\[
= \frac{1}{N_G^2} \sum_{G,G'} u_{kp}^{*}(G) u_{kq}^{*}(G') \sum_A \sum_{G''} V_{A}^{en}(G - G') e^{-i(G - G')r_A}. 
\]

While there are no approximations in this result other than the discreteness of the Fourier grid, it is not particularly amenable for computation due to the scaling of \( O(N_G^2) \). One way to reduce the cost of this is instead of rewriting the individual orbitals \( \phi_{kp}(r) \) and \( \phi_{kq}(r) \) in terms of their expansion of planewaves, we can instead expand their \emph{product} in terms of planewaves. It should be obvious that expanding each over the same incomplete basis these two should give different answers \footnote{One can think of an easy case for planewaves where the two answers will obviously differ: one can expand each orbital in terms of auxiliary planewaves with maximum frequency component \( |G| \). The product of each of these orbitals will then have a maximum frequency component of \( 2|G| \) since frequency components add when the planewaves are multiplied together. Looking at the product of these two orbitals and \emph{then} expanding in terms of planewaves, we see the answer will be quite different since now our expansion will contain terms only up to a frequency component \( |G| \).}

In the context of a planewave basis, the choice of expanding individual orbitals and products of orbitals can lead to quite different answers both numerically and in terms
of scaling. We will show now where the difference lies and when the two expansions
will be identical. We can first expand the individual orbitals over the planewaves and
obtain

\[ \psi_{pkp}(r) = \sum_{G,G'} u_{pkp}(G) u_{qkp}(G') e^{i(G'-G)r} \]

Beyond this there are no simplifications unless one can assume an infinite (complete)
set of reciprocal vectors \( G \) and \( G' \). In the limit of an infinite number of reciprocal
vectors, we have that the summation of \( (G' - G') \) is just the set \( G \) and so this reduces
to

\[ \rho_{pkp,qkp}(r) = \psi_{pkp}^*(r) \psi_{qkp}(r) \]

where we have

\[ \rho_{pkp,qkp}(r) = \psi_{pkp}^*(r) \psi_{qkp}(r) \]
\[ \rho_{pkp,qkq}(G) = \int dr \ e^{-i(k_q-k_p+G)r} u^*_p(r) u_q(r). \]  

(87)

Thus we have shown that the product of the expansion over planewaves is equal to the expansion of the product of planewaves in the limit of a complete set of reciprocal vectors. We note that in doing so, we have reduced the original scaling for the resolution over planewaves of \( \mathcal{O}(N^2_a) \) to a scaling of \( \mathcal{O}(N_G) \). For the most part, the cost of the scaling outweighs any additional error created from this approximation and so you will find most codes choose to expand the density.

Using a resolution of the identity for the density, we have that the electron-nuclear potential can be written as

\[ V_{k_pq} = \int dr \ \rho_{pkp,qkq}(r)V(r) \]

\[ = \frac{1}{V N^2_G} \int dG \ \sum_{G'} \rho_{pkp,qkq}(G) \sum_{G'} V(G') e^{i(G+G')r} \]

\[ = \frac{1}{V N^2_G} \sum_{G,G'} \rho_{pkp,qkq}(G) V(G') \int dr \ e^{i(G+G')r} \]

\[ = \frac{1}{N^2_G} \sum_{G,G'} \rho_{pkp,qkq}(G) V(G') \delta_{G,-G'} \]

\[ = \frac{1}{N^2_G} \sum_{G} \rho_{pkp,qkq}(G) V(-G). \]  

(88)

As an aside we note that there is both the \( G \) and \( -G \) as the Fourier components of the above which will add to zero. A useful check for any of these matrix elements between states of zero momentum is to make sure that the matrix elements also add up to zero momentum. Technically speaking, we could also have written this with \( V(G) \) instead of \( V(-G) \) since these are equal but it helps to keep minus signs for purposes of checking.

Note however that for the density summation our \( G \) here technically is the sum over \( (G - G') \), which again arises since the momentum is additive when multiplying the two Fourier components of the wavefunction used in creating the density. This usually means that the density will be done on a larger grid to take into account these Fourier components with larger momenta. However, there is nothing stopping one from
using the same density grid as the grid for the wavefunction, and this is how the density grid is treated within the periodic module of PySCF [10]. While one finds generally a slightly slower convergence with respect to the basis set, there is the benefit of reduced bookkeeping due to using only one grid and a modest increase in speed.

4.2.4 Electron Repulsion Integrals

We now move on to the most computationally challenging part of the problem in solid state systems. While the analytical evaluation of the electron repulsion integrals is relatively straightforward, the problem arises in the speed at which they can be evaluated. For large systems, there are too many integrals to be held in memory and so they would need to either be held on disk or be built as needed (on-the-fly). The on-the-fly method is certainly the most scalable method as it requires no storage on disk. The various methods to speed the integrals up will not be discussed here, but it is just helpful to know some of the current limitations and ways to improve methods for the solid state.

For the electron repulsion integral written in physicist’s notation, \( \langle \phi_{p_{k_p}} \phi_{q_{k_q}} | \phi_{r_{k_r}} \phi_{s_{k_s}} \rangle \), we have
\[ \langle \phi_p \phi_q, \phi_{r, s} \rangle = \int \! \! dr_1 dr_2 \, \phi_p^* (r_1) \phi_{r, s} (r_1) \frac{1}{|r_1 - r_2|} \phi_q^* (r_2) \phi_{s, k_s} (r_2) \]

\[ = \int \! \! dr_1 dr_2 \, \rho_{p, r, s} (r_1) \frac{1}{|r_1 - r_2|} \rho_{q, k_q, s} (r_2) \]

\[ = \frac{1}{N_G^2} \int \! \! dr_1 dr_2 \sum_G \rho_{p, r, k_r} (G) e^{i(k_r - k_p + G) r_1} \frac{1}{|r_1 - r_2|} \]

\[ \times \sum_{G'} \rho_{q, k_q, s} (G') e^{i(k_q - k_s + G') r_2} \]

\[ = \frac{1}{N_G^2} \int \! \! dr_1 dr_2 \sum_G \rho_{p, r, k_r} (G) e^{i(k_r - k_p + G - k_s + k_q - G') r_1} \]

\[ \times \frac{1}{|r_1 - r_2|} \sum_{G'} \rho_{q, k_q, s} (G') e^{i(k_q - k_s + G') (r_2 - r_1)} \]

\[ = \frac{1}{N_G^2} \sum_{G, G'} \rho_{p, r, k_r} (G) \rho_{q, k_q, s} (G') \int \! \! dr_1 \, e^{i(k_r - k_p + G - k_s + k_q - G') r_1} \]

\[ \times \int \! \! dr_2 \, \frac{1}{|r_1 - r_2|} \, e^{i(k_q - k_s + G') (r_2 - r_1)} \]

(89)

where in going from the third to fourth line we inserted \( e^{-i(k_q - k_s + G') r_1} e^{i(k_q - k_s + G') r_1} \).

Now we focus on the inner integral and make the change of variables \( r = r_2 - r_1 \), and notice that this is simply the Fourier component of the potential operator

\[ \int \frac{1}{|r|} e^{-i(k_q - k_s + G') r} = V(k_q - k_s + G'). \] (90)

And so our earlier equation reduces to

\[ = \frac{1}{N_G^2} \sum_{G, G'} \rho_{p, r, k_r} (G) \rho_{q, k_q, s} (G') \int \! \! dr_1 \, e^{i(k_r - k_p + G - k_s + k_q - G') r_1} \]

\[ \times V(k_q - k_s + G') \]

\[ = \frac{V}{N_G^2} \sum_{G, G'} \rho_{p, r, k_r} (G) \rho_{q, k_q, s} (G') \delta_{G', k_q - k_p + k_r - k_s + G} \]

\[ \times V(k_q - k_s + G') \]

\[ = \frac{V}{N_G^2} \sum_G \rho_{p, r, k_r} (G) \rho_{q, k_q, s} (k_q - k_p + k_r - k_s + G) V(k_q - k_s + G') \] (91)

We see that the momentum conservation occurs naturally in the derivation through
the delta function term $\delta_{G', k_q - k_p + k_r - k_s + G}$, and that this delta function is nothing more than the earlier stated conservation condition in equation 77. Here the issue of aliasing makes its appearance, where we find that the momentum argument $k_s - k_q + G'$ for the potential can fall outside of our original set of planewaves. While this in itself isn’t a problem, there are some subtleties that arise when getting agreements between supercell calculations and k-point calculations. As stated before, doing for example a $3 \times 3 \times 3$ k-point calculation should agree with a large $3 \times 3 \times 3$ supercell done at the gamma point and with a concomitant basis set. Without proper aliasing, one will have small differences between the calculations if done at a finite sampling of planewaves. Being able to check the whether the supercell calculation and k-point calculation agree at a finite (small) number of planewaves is a quick and easy way to check the correctness of any periodic implementation.

More importantly, without aliasing the integrals will not have the correct permutation symmetry for a finite sampling of planewaves. We can see that for an integral $\langle \phi_p, \phi_q | \phi_r, \phi_s \rangle$ we would normally have the symmetries

$$
\langle \phi_p, \phi_q | \phi_r, \phi_s \rangle = \langle \phi_q, \phi_p | \phi_s, \phi_r \rangle = \langle \phi_s, \phi_r | \phi_p, \phi_q \rangle = \langle \phi_r, \phi_s | \phi_q, \phi_p \rangle.
$$

However, we see that for the first and fourth symmetries, we would have two different arguments for the potential term in 91. Letting $q = k_s - k_q$ and $q' = k_r - k_p = -q + G'$, we can see this combination of k-points satisfies the momentum conservation condition. The first integral symmetry above contains the term $V(q + G)$ while the second contains $V(-q + G + G')$. For a finite set of planewaves we will not necessarily have equality between these supposedly symmetry quantities. For an infinite set of planewaves, however, we see that we can rewrite the set $\{G + G'\}$ as $\{G\}$, and note that we then have
\[
V(-q + G + G') = V(-q + G)
\]
\[
= V(-q - G)
\]
\[
= V(q + G),
\]

where we used the fact that for an odd number of grid points \{-G\} = \{G\} and so again we would regain the correct permutational symmetries.

To correct this issue of aliasing, we merely move high frequency components of \(q + G\) down to its lower frequency counterpart. Mathematically, for a set of \(G\) vectors as described in equation 49 where \(|m_i| \leq \bar{m}_i\), we can describe any element \(q\) as

\[
q = \sum_{i=1,2,3} b_i (\bar{m}_i + 0.5) G_i,
\]

where \(|b_i| \leq 1\) and \((\bar{m}_i + 0.5)\) represents the boundary of the “G-cube”. Any momentum vectors originally outside this boundary will be translated by finding the \(b_i\) for its corresponding momentum within the G-cube.
5 Spectral Functions of the Uniform Electron Gas via Coupled-Cluster Theory and Comparison to the GW and Related Approximations.

Having discussed the coupled-cluster theory and the basics of periodic systems, it’s now possible to apply the equation of motion theory to compute the excited states of periodic systems. Here we focus on the uniform electron gas due to its ease of theory, as the Fock operator is diagonalized in the space of planewaves, and for the difficulties that this metallic system gives for perturbation theories. In fact, Shepherd and coworkers [48] were the first to show that the MP2 energy diverges with increasing system size for the finite jellium model. In our work it has been shown that the coupled-cluster performs well when compared against the GW method that is the standard for computing excited states in periodic systems. While the calculations here did not make use of the full symmetry available in these systems, such as momentum symmetry, we were still able to run the coupled-cluster on moderate sized systems and is a viable
option for excited states in periodic systems.

5.1 Introduction

Computing the electronic excitations and spectra of condensed phase systems with significant correlations from first-principles continues to be a premier challenge in computational materials science. Currently, a widely used approach is time-dependent many-body perturbation theory (MBPT). In this approach, the electronic Green’s function $G$, whose poles yield the single-particle excitation energies, is obtained by evaluating Feynman diagrams representing many-electron interaction processes. Retaining only the lowest-order diagram in an expansion in terms of the screened Coulomb interaction $W$ leads to the $GW$ method [49]. The $GW$ method greatly improves band gaps obtained from density-functional theory (DFT) [14, 15], and further yields other accurate quasiparticle properties, such as lifetimes and bandwidths [50, 51], in a wide range of weakly and moderately correlated materials.

However, despite its successes, the $GW$ method has well-known limitations. Specifically, it has proven difficult to systematically improve $GW$ theory by including higher-order Feynman diagrams, so-called vertex corrections. While extensions of the $GW$ approach have been developed for specific applications – such as the cumulant expansion of the time-dependent Green’s functions for the description of plasmon satellites [18, 52, 53] or the $T$-matrix approach for magnetic systems [54–56] – there exists currently no universally accepted and applicable “beyond-$GW$” approach. An additional problem in most practical $GW$ calculations is a dependence of the results on the mean-field starting point. This arises because most implementations apply the $GW$ method as a perturbative “one-shot” correction to a mean-field calculation, such as DFT or Hartree-Fock (HF); this is usually referred to as the $G_0W_0$ approach. At a greater numerical cost, self-consistent $GW$ calculations have been carried out with mixed success [57–61].

More common in $ab$ $initio$ quantum chemistry, methods based on time-independent many-body perturbation theory provide a different route to electronic excitations [28, 62–64]. In this framework, coupled cluster theory is an example of a well-studied and

\[ \text{iii} \text{This work is based primarily off our previously published work [21].} \]
systematically improvable hierarchy within which to resum the corresponding classes of Goldstone diagrams \[28, 65, 66\]. Electronic excited states are obtained by equation-of-motion (EOM) coupled-cluster theory \[32, 67, 68\]. For molecules with weak to moderate correlations, coupled-cluster theories at the singles, doubles, and perturbative triples level are established as the quantitative “gold standard” of quantum chemistry \[66\].

While such \textit{ab initio} coupled-cluster theories have been widely applied to atoms and molecules, they have traditionally been thought too expensive to use in extended systems; for example, coupled-cluster theory with single and double excitations formally has a computational scaling $O(N^6)$. However, with improvements in algorithms and increases in computer power, the exciting possibility of applying these methods to condensed matter problems is now within reach. For example, very recent work has applied \textit{ground-state} coupled-cluster theory to the uniform electron gas (UEG) \[48, 69, 70\] as well as atomistic solids \[8\]. Correlated \textit{excited states} are the next frontier.

In this Letter, we apply, for the first time, EOM coupled-cluster theory to the UEG and study its one-particle electronic excitations. The UEG is a paradigmatic model of metallic condensed matter systems and these calculations illustrate the potential of applying coupled-cluster theory in first-principles materials simulations. We employ coupled-cluster theory with single and double (and in some cases triple) excitations; at this level, the diagrammatic content of our treatment goes significantly beyond the standard \textit{GW} level of approximation. As such, our coupled-cluster spectra allow us to assess the quality of vertex corrections to the \textit{GW} method in the UEG. For example, we evaluate the accuracy of the \textit{GW}-plus-cumulant treatment of the correlated satellite structure. Further, as we consider the electron gas at the density of $r_s = 4.0$ corresponding to that of elemental sodium, our results for the occupied bandwidth provide strong evidence to settle the long-standing puzzle concerning the interpretation of photoemission experiments in this material.
5.2 Methods

We study electronic excitations of the three-dimensional UEG using a supercell approach, i.e. we place \( N \) electrons in a cubic box of volume \( \Omega \) with a neutralizing positive background charge and periodic boundary conditions. The thermodynamic limit is obtained, in principle, by increasing \( N \) and \( \Omega \) while keeping the density \( N/\Omega \) fixed. Here, we only present results for the UEG with a Wigner-Seitz radius \( r_s = 4.0 \) (\( k_F = 0.480 \) a.u.) corresponding approximately to the valence electron density of metallic sodium.

For the UEG Hamiltonian \( \text{iv} \) we calculate the one-electron Green’s function \( G_k(\omega) \) and the corresponding spectral function \( A_k(\omega) = \pi^{-1} |\text{Im} G_k(\omega)| \) using several methods: (i) mean-field theory, i.e. HF and DFT in the local-density approximation (LDA) [74], (ii) time-dependent MBPT, i.e. the \( GW \) and \( GW+C \) methods, (iii) EOM coupled-cluster theory, and (iv) dynamical density matrix renormalization group (DMRG), which provides numerically exact spectral functions for small system sizes [75]; all DMRG calculations were performed with a bond dimension of \( M = 1000 \). Specifically, we compute spectral functions of occupied states, which are the ones probed in photoemission experiments.

The one-particle eigenstates of the mean-field theories are plane-waves, \( \phi_k(r) = \Omega^{-1/2} e^{i k \cdot r} \). These serve as a finite basis set, with a cutoff \( k_{\text{cut}} \), in the subsequent MBPT, CC, and DMRG calculations. The corresponding eigenenergies are given by \( \epsilon_k = k^2/2 + V_{\text{xc}}^k \), where \( V_{\text{xc}}^k \) denotes the exchange-correlation matrix element, evaluated either at the HF or DFT-LDA level (the Hartree term exactly cancels the interaction energy with the positive background charge density).

Based on the HF and DFT-LDA mean-field starting points, we carry out one-shot \( GW \) (i.e. \( G_0 W_0 \)) calculations [14, 15] where screening is treated in the random-phase approximation, as well as \( G_0 W_{\text{xc}} \) calculations where screening is treated with the DFT-LDA dielectric function [56, 76]. We also evaluate spectral functions using the \( GW \)+plus-cumulant (henceforth \( GW+C \)) method. This approximation yields the

\[ iv \text{We treat the divergent } G = 0 \text{ component of the Coulomb potential with the “probe-charge” Ewald summation method [71], i.e. } v_{G=0} = a_0/L \text{ where } a_0 = 2.837 \text{ is the Madelung constant of a 3D simple cubic lattice [72, 73].} \]
exact solution for a dispersionless core electron interacting with plasmons [77] and noticeably improves the description of plasmon satellite properties compared to \( GW \), while retaining the accuracy of \( GW \) for the quasiparticle energies. The \( GW+C \) formalism defines the Green’s function as

\[
G_k(t) = G_{0,k}(t) \exp \left[ -i \Sigma_k(t) + C_k(t) \right],
\]

where \( G_0 \) is the Green’s function from mean-field theory, \( \Sigma_k \) is the bare exchange self-energy and \( C_k(t) = \pi^{-1} \int d\omega |\text{Im}\Sigma_k(\omega + E_k^{GW})(e^{-i\omega t} + i\omega t - 1)/\omega^2 | \) is the cumulant function [52, 53, 78]. Here, \( E_k^{GW} \) denotes the \( GW \) orbital energy. The \( GW+C \) approach has been applied to range of bulk materials [18, 19, 79, 80] and nanosystems [20, 81] and good agreement with experimental measurements on satellite structures was found. However, comparisons of the \( GW+C \) to other accurate numerical calculations have been difficult to perform, and this is one of the objectives below.

We perform EOM coupled-cluster calculations of the one-electron Green’s function starting from the mean-field ground-state determinant \( |\Phi_0\rangle \), defined by the occupied one-particle eigenstates with \( k < k_F \). We briefly describe the relevant theory below; we refer to Ref. [28] for details. The coupled-cluster ground-state is defined as

\[
|\Psi_0\rangle = e^T|\Phi_0\rangle,
\]

where the cluster operator is

\[
T = \sum_{ia} t_{ia} c_a^\dagger c_i + \frac{1}{4} \sum_{ijab} \frac{t_{ijab} c_a^\dagger c_b^\dagger c_j c_i}{2} + \ldots
\]

(with the indices \( i, j \) referring to occupied states and the indices \( a, b \) referring to unoccupied states). Singles, doubles, and triples coupled-cluster theories (denoted CCS, CCSD, and CCSDT) correspond to truncating \( T \) after one, two, and three electron-hole excitations. The \( T \) operator and coupled-cluster ground-state energy are obtained through the relations

\[
E_0 = \langle \Phi_0 | e^{-T} \hat{H} e^T | \Phi_0 \rangle = \langle \Phi_0 | \hat{H} | \Phi_0 \rangle = \langle \Phi_0 | \hat{H} | \Phi_0 \rangle = \ldots \tag{95}
\]

where the notation \( \Phi_0^a, \Phi_0^{ab}, \ldots \) represents Slater determinants with one, two, \ldots electron-hole pairs, and \( \hat{H} \) is the non-Hermitian coupled-cluster effective Hamiltonian. By construction from Eq. (95), \( |\Phi_0\rangle \) is the right ground-state eigenvector of \( \hat{H} \); its left ground-state eigenvector \( \langle \tilde{\Phi}_0 | \) takes the form \( \langle \tilde{\Phi}_0 | (1 + S) \), where \( S = \sum_{ia} s_i c_a^\dagger c_i^\dagger + \frac{1}{4} \sum_{ijab} s_{ij} c_a^\dagger c_b^\dagger c_j c_i + \ldots \) creates excitations in the bra, to the same level as in \( T \).
Coupled-cluster excited states and energies are formally determined by diagonalizing \( \hat{H} \) in an appropriate space of excitations. For the single-particle (ionization) energies here, we diagonalize in the space of 1-hole \((1h)\) and 2-hole, 1-particle \((2h1p)\) states for a CCSD ground-state, additionally including the space of 3-hole, 2-particle \((3h2p)\) states for a CCSDT ground-state [82, 83]. The ionization contribution to the CC Green’s function [22, 23] is then defined in the same space, as

\[
G_k(\omega) = \langle \Phi_0 | c_k^\dagger P \frac{1}{\omega - (E_0 - H) - i\eta} P c_k | \Phi_0 \rangle
\]  

(96)

where \( P \) projects onto the space of \( 1h, 2h1p, \) and (for CCSDT) \( 3h2p \) states. We emphasize that although the initial ground-state CCSD calculation scales as \( O(N^6) \), the excited state ionization-potential EOM-CCSD has a reduced scaling \( O(N^5) \); this should be compared to the \( O(N^4) \) scaling of \( GW \) methods.

With respect to other works, this Letter represents multiple significant methodological advances. Most importantly, we present the first application of CCSD to the full spectrum of excited states for a condensed phase system and establish its accuracy in a parameter regime relevant for real materials. These results complement recent work applying CCSD to the ground state of the electron gas [48, 69, 70]. Remarkably, to the best of our knowledge, our results are also the first report of the full frequency-dependent CCSD spectral function (and not just the energy of select ionization poles), for any system. Furthermore, we present the first nonperturbative CCSDT results for the ground state of the UEG, as well as the first CCSDT Green’s function for any system.

### 5.3 Analysis of CC and GW methods

Coupled-cluster theory with \( n \)-fold electron-hole excitations in the \( T \) operator includes all time-independent diagrams with energy denominators that sum at most \( n \) single-particle energies. At the singles and doubles CCSD level (the lowest level used in this work), this already includes more Feynman diagrams than are in \( GW \) theory. In particular, the CCSD energies and Green’s function include not only the ring diagrams
which dominate the high-density limit of the electron gas [84] and which yield the screened RPA interaction in GW, but also ladder diagrams (such as generated in T-matrix approximations) and self-energy insertions which couple the two [85].

Unlike GW theory, CC approximations are invariant to the values of the single-particle energies in the mean-field used to generate $|\Phi_0\rangle$. They are further relatively insensitive to the single-particle orbitals, because $e^{T_1}$ parametrizes rotations from $|\Phi_0\rangle$ to any other determinant [86]. While CC calculations typically start from a HF mean-field calculation, in the UEG the HF and DFT mean-field theories share the same plane-wave states as their one-particle eigenstates. This means that the UEG CC calculations are completely invariant to the mean-field choice (in the paramagnetic phase). This complicates a fair comparison between one-shot GW calculations and the CC calculations. For this reason, we present calculations with both HF (HF+GW and LDA (LDA+GW) as a reference; the former may be considered a fairer comparison with CC when assessing the diagrammatic quality of the theories.

5.4 Results

To establish the accuracy of the different methods, we initially study a supercell containing 14 electrons in a minimal single-particle basis of 19 spatial orbitals ($k_{\text{cut}} = 0.572$ a.u.). The electrons occupy seven orbitals, namely the orbital with $k = (0, 0, 0)$, corresponding to the bottom of the band in the thermodynamic limit, and the six-fold degenerate highest occupied orbital $k = (2\pi/L, 0, 0)$ corresponding to the Fermi level in the thermodynamic limit. For this small system, we can compare GW and CCSD to coupled-cluster theory with all triple excitations (CCSDT) as well as numerically exact dynamical density matrix renormalization group (DMRG) calculations of the spectral function.

Figure 1(a) shows our results for the deeply bound $k = (0, 0, 0)$ state. All spectral functions (except for $GW+C$) exhibit two peaks: a quasiparticle peak near $-6$ eV and a strong satellite peak near $-10$ eV. We find excellent agreement between the CCSDT and the dynamical DMRG result. The agreement between CCSD and the DMRG
result is also very good, in particular for the quasiparticle peak. Starting from the same HF reference as typically used in coupled-cluster theory, HF+GW yields a much less accurate result: the binding energy of the quasiparticle is too large by about 1 eV and the spectral weight is overestimated by almost a factor of 2. This error is inherited from the underlying HF mean-field theory and illustrates the starting point dependence of the method. Even worse results are obtained for the satellite feature which is at far too low an energy. However, when starting from a DFT-LDA reference, the GW approximation gives results with much improved accuracy, and is only slightly worse than CCSD. (As discussed above, in the UEG the CC results are invariant to the reference).

Interestingly, GW+C yields several satellite peaks with incorrect energies and underestimated peak heights, illustrating some of the challenges in systematically improving on GW theory through standard vertex corrections. By construction, the GW+C approach produces a plasmon-replica satellite structure (see below) even for small systems, which is physically incorrect.

Consistent with Fermi liquid theory, the spectral functions of the $\mathbf{k} = (2\pi/L, 0, 0)$ state shown in Fig. 1(b) exhibit significantly weaker electron correlations than the spectral functions of the $\mathbf{k} = (0, 0, 0)$ state. Specifically, all methods predict a strong quasiparticle peak with a binding energy of about 5 eV and weak satellite features. The inset of Fig. 1(b) shows that the detailed structure of the satellites is quite complex. While CCSDT accurately captures the complex features seen in the exact spectrum, none of the other methods are fully satisfactory. In particular, HF+GW pushes satellite features to too low energies, the LDA+GW places the satellite peaks at too high an energy, and CCSD places them in between. GW+C correctly reduces the weight of the main GW satellite peaks but does not otherwise improve the spectrum.

Next, to study the approach to the thermodynamic limit, we carried out calculations on larger supercells for which CCSDT and dynamical DMRG are no longer computationally tractable. We performed CCSD, GW, and GW+C calculations for supercells containing 38, 54, 66, and 114 electrons. The quasiparticle features of all systems
Figure 1: Spectral functions for the UEG with \( r_s = 4.0 \) using a supercell containing 14 electrons in 19 spatial orbitals. (a) For the \( k = (0, 0, 0) \) state, the spectral functions exhibits a prominent satellite peak; the HF+GW result has been scaled down by a factor of 1.5. (b) For the highest occupied state at \( k = (2\pi/L, 0, 0) \), the spectral function exhibits a strong quasiparticle peak with a very weak satellite structure; the satellite region between \(-18 \text{ eV} \) and \(-7 \text{ eV} \) has been magnified by factor of 30. A linewidth broadening of \( \eta = 0.2 \text{ eV} \) is used in all calculations.

studied are similar (e.g. the occupied bandwidth), however the satellite features are unsurprisingly different, and so here we will only discuss the largest system studied.

For the 114 electron system, we used plane-wave basis sets with at least 485 spatial orbitals, corresponding to \( k_{\text{cut}} = 0.985 \text{ a.u.} \), which is sufficiently large to converge all peak positions to within 0.2 eV.

Figure 2(a) shows the spectral function of the \( k = (0, 0, 0) \) state for the UEG with 114 electrons in 485 orbitals. The CCSD spectral function exhibits a strong quasiparticle peak near \(-6 \text{ eV} \). For the GW calculations, we observe again a strong dependence on the mean-field starting point: while the quasiparticle energy from LDA+GW agrees very well with CCSD, that from HF+GW is significantly worse. This is not surprising since DFT-LDA yields much more accurate metallic bands than HF.

At higher binding energies, the CCSD spectral function exhibits a rather complex
satellite structure, however two major regions of spectral weight can be identified near
12 eV and 18 eV. In contrast, both the HF+GW and the LDA+GW spectral functions exhibit only a single, prominent satellite peak. Lundqvist and co-authors [87, 88] assigned this peak to a novel excited state, the plasmaron, which is a coherent superposition of strongly coupled plasmon-hole pairs. While several experiments reported the observation of plasmaron excitations in doped graphene and semiconductor quantum-well two-dimensional electron gases, it has recently become clear that their prediction by GW is spurious. Vertex-corrected time-dependent MBPT approaches, such as the GW+C method, do not predict such a state and instead yield a satellite structure that consists of an infinite series of peaks corresponding to the “shake-up” of one or more plasmons [52, 77]. Notably, the major peaks in the CCSD spectral function are separated by roughly 6 eV corresponding to the classical plasma frequency $\omega_p = 5.9$ eV in an electron gas with $r_s = 4.0$. Comparing the LDA+GW+C result to CCSD in Fig. 2(a), we find a qualitatively similar spectrum. However, at least at this system size, the CCSD spectral function has a stronger quasiparticle peak, a larger spectral width, and significantly more fine-structure than the GW+C spectral function.

To assess remaining errors of the 114 electron system relative to the thermodynamic limit, we compare the $k = (0, 0, 0)$ spectral functions of the UEG with 114 electrons with the results fully converged to the thermodynamic limit for the LDA+GW and the LDA+GW+C methods. Fig. 2(b) shows good qualitative agreement between the two sets of spectral functions for this class of methods.

Finally, Fig. 2(c) shows the quasiparticle energies as function of the electron wave vector, i.e. the energy dispersion relation, for the 114 electron system \footnote{For this data, we performed an extrapolation to the complete basis set (CBS) limit. For each calculation with $M$ basis functions, the quasiparticle peaks were fitted with a Lorentzian lineshape. These peak positions were observed to have a $1/M$ dependence, allowing for extrapolation to $M \rightarrow \infty$. As a function of wavevector, we then fitted the CBS-limit peak positions to a quadratic “effective mass” dispersion and referenced all energies to the extrapolated Fermi energy in the thermodynamic limit.}. The inferred bandwidths are 2.96 eV for CCSD, 3.79 eV for HF+GW, 2.77 eV for LDA+GW, and 2.56 eV for LDA+GW$_{xc}$; self-consistency treated within the quasiparticle self-consistent GW scheme gives only a minor bandwidth narrowing compared to LDA+G$_0$W$_0$ [60].
While DFT-LDA gives a bandwidth of 3.13 eV, HF predicts a value of 7.29 eV, significantly larger than any other method. The failure of HF to describe metallic systems is well-documented and results from the absence of screening.

The bandwidth of simple metals, and in particular sodium, has been the subject of a decades-long debate. Plummer and co-workers [89, 90] carried out angle-resolved photoemission experiments on sodium and reported a bandwidth of 2.5–2.65 eV, significantly smaller than the free-electron and DFT-LDA value of ~3.1 eV, and even the LDA+GW value of ~2.8 eV [49]. Interestingly, the experimental result agrees quite well with the bandwidth from a LDA+GW_{xc} calculation [56, 76], which contains vertex corrections for the dielectric function; however, including vertex corrections also in the self-energy increases the bandwidth again [91–93]. As an alternative explanation, Shung and Mahan [94, 95] suggested that the measured bandwidth results from many-body effects in combination with final-state effects and an interference between surface and bulk photoemission. The close agreement seen here between the quasiparticle dispersion of LDA+GW and CCSD – especially the larger bandwidth of CCSD – suggests that the theoretical description of the quasiparticle peak positions may be adequate already and supports Shung and Mahan’s thesis that the remaining discrepancy in the observed bandwidth is due to final-state and interference effects.

### 5.5 Conclusion

We have demonstrated the first application of coupled-cluster techniques to the computation of spectra in condensed phase systems, using the uniform electron gas as a model system. For finite uniform electron gas models of various sizes we find that coupled-cluster, even at the singles and doubles level (CCSD), provides improvement over GW and even GW-plus-cumulant theory. Interestingly, while the latter exhibits good accuracy for large systems (producing reasonable plasmon-like satellite structures), the former is significantly more accurate for small systems; CCSD naturally interpolates between these two limits. In conclusion, by providing a systematic framework that goes beyond the diagrammatic content of the GW approximation, coupled-cluster theories
represent a very promising, new direction in the search for more accurate methods to compute the spectra of real materials.
Figure 2: (a) Spectral function of the \( k = (0, 0, 0) \) state of the 3D UEG with \( r_s = 4.0 \) and 114 electrons in 485 orbitals. The HF+GW result is scaled down by a factor of 1.5 and a linewidth broadening of \( \eta = 0.8 \) eV is used in all calculations. (b) Comparison of the spectral functions of the \( k = (0, 0, 0) \) state in the thermodynamic limit (solid curves) and the 114 electron system (dashed curves) from LDA+GW (blue curves) and LDA+GW+C (green curves). (c) Complete basis set limit quasiparticle energies as a function of wave vector for the 114 electron system (symbols) and quadratic fits (dashed curves).
6 Gaussian-based Coupled-Cluster Theory for the Ground State and Band Structure of Solids

This last section will cover our recent work applying these coupled-cluster methods to “real” systems. While our previous work did not use any of the periodic symmetries, here the symmetry was exploited to run rather large calculations with over 2000 orbitals. Here we were interested in the calculation of the band gap of diamond and silicon, two relatively easy (from a correlation standpoint) and common systems where there is still error on the order of tenths of an eV for methods such as GW and its variants. These calculations represent the first of their kind in three-dimensional periodic systems vi.

6.1 Introduction

The electronic ground-state and excited-state structure of three-dimensional solids and low-dimensional nanomaterials is an ongoing challenge in computational materials science. A theoretical program based on density functional theory (DFT), frequently

viThis work is based primarily off our previously published work [McClain2017 ].
combined with time-dependent many-body perturbation theory, has become a reliable and accurate standard approach. In particular, ground-state properties of weakly correlated materials are well-described by DFT, the principal charged excited states leading to band structure can be treated with hybrid [96–98] or semilocal [99] functionals, as well as the many-body GW approximation [14, 15], and neutral excited states can be described with time-dependent DFT [100, 101] or the Bethe-Salpeter equation [102]. However, errors in this approach can be difficult to assess, due to (for example) the underlying DFT exchange-correlation functional or self-consistency in the GW approximation. Even in the formal many-body frameworks, systematic improvements (vertex corrections) are only recently being pursued, with mixed success [103–106].

In contrast, wavefunction-based approaches originating in the quantum chemistry community are now starting to be applied to the properties of condensed-phase materials. Atomistic one- and two-dimensional systems have been studied by configuration interaction with single excitations (CIS), time-dependent Hartree-Fock (HF), and time-dependent DFT [107], as well as second-order Møller-Plesset perturbation theory (MP2) [11, 12, 108, 109] and coupled-cluster (CC) theory with single and double excitations (CCSD) [11, 12, 108]. For three-dimensional solids, MP2 has been implemented with local correlation in the CRYSCOR program [110, 111], in the plane-wave based VASP package with full Brillouin zone sampling [112–114], and in the mixed Gaussian and plane-wave CP2K package at the Γ-point [115]. Most recently, ground-state CCSD, CCSD with perturbative triple excitations (CCSD(T)), and full CI quantum Monte Carlo for periodic solids have been reported [8, 114]. We also mention the parallel development of non-periodic cluster-based approaches for solids, namely incremental methods [116–119] and embedding techniques [120–122].

Recently, we performed the first excited-state CC treatment of a three-dimensional condensed-phase system, at the equation-of-motion (EOM)-CCSD and EOM-CCSDT level, focusing on the uniform electron gas at the density of metallic sodium [21]. For small unit cells in a finite single-particle basis, the EOM-CCSDT one-particle spectral function was in nearly perfect agreement with numerically exact (time-dependent den-
sity matrix renormalization group) results, and more accurate than any tested flavor of the \(GW\) approximation, while the EOM-CCSD one-particle spectral function quantitatively improved on the standard \(GW\) approximation. Importantly, unlike the \(GW\) approximation (in practice, \(G_0W_0\)), the EOM-CC spectra are invariant to the starting choice of single-particle energies. At larger system sizes approaching the thermodynamic limit, it was not possible to carry out full EOM-CCSDT calculations, but it was still possible to compute the EOM-CCSD spectrum. Similarly to the \(GW\)-plus-cumulant approximation \([18, 52]\), the EOM-CCSD spectrum removed the fictitious plasmaron pole of the \(GW\) spectrum \([87, 88]\), demonstrating that it correctly treats the coupling of charged excitations to bosonic plasmon excitations in metals.

Here, we continue this program and present the first results of CC excited states for atomistic solids. The layout of this paper is as follows. In Sec. 6.2, we describe the Gaussian basis sets, integral evaluation, and HF calculations for periodic systems, including issues associated with divergent terms and finite-size effects. In Sec. 6.3, we describe the ground-state CCSD and excited-state EOM-CCSD formalisms for periodic systems, as well as the connection to quasiparticle band structure. In Sec. 6.4 we discuss the convergence and finite-size effects of our HF and correlated calculations. We present ground-state HF, MP2, and CCSD results in Sec. 6.5 for diamond and silicon. This section also includes the technical details of our calculations, a discussion of convergence, and comparison to existing MP2 and CCSD calculations on solids \([8, 113]\). In Sec. 6.6, we present EOM-CCSD band structures and a convergence study of the indirect band gap. We conclude in Sec. 6.7.

6.2 Gaussian-based periodic integrals and Hartree-Fock theory

We use an underlying single-particle basis of crystalline Gaussian-based atomic orbitals (AOs). These are translational-symmetry-adapted linear combinations of Gaussian AOs of the form

\[
\phi_{\mu k}(r) = \sum_T e^{i k \cdot T} \tilde{\phi}_{\mu}(r - T) = e^{i k \cdot r} u_{\mu k}(r)
\]  

(97)
where $T$ is a lattice translation vector and $k$ is a crystal momentum vector in the first Brillouin zone. In this work, we sample $k$ from a uniform (but not necessarily Monkhorst-Pack) grid. The second equality above expresses Bloch’s theorem, where the Bloch function $u_{\mu k}(r)$ is fully periodic with respect to all lattice translations. Therefore, we can exactly expand the crystalline AOs in a set of auxiliary plane-waves

$$
\phi_{\mu k}(r) = \sum_G \phi_{\mu k}(G)e^{i(k+G) \cdot r},
$$

$$
\phi_{\mu k}(G) = \frac{1}{\Omega} \int_\Omega d\tau \phi_{\mu k}(r)e^{-i(k+G) \cdot r},
$$

where $G$ is a reciprocal lattice vector and $\Omega$ is the unit cell volume.

Given this basis choice, there are several different ways to evaluate the corresponding overlap, kinetic energy, electron-nuclear attraction, and Coulomb integrals arising in quantum chemistry. We will present a careful comparison of different choices in a later publication, and here only detail the grid-based scheme that we have used for the results in this work, which closely resembles the Gaussian and plane-wave scheme used in the CP2K package [1, 2]. We use a dual real-space and reciprocal-space representation of the crystalline AOs. In real space, we represent $\phi_{\mu k}(r)$ on a uniform real-space grid in the unit cell. In reciprocal space, we obtain $\phi_{\mu k}(G)$ by a fast Fourier transform (FFT) leading to a representation on a uniform grid of points corresponding to the reciprocal lattice vectors of our unit cell. The real-space grid density roughly corresponds to a kinetic energy cutoff in the plane-wave representation. To ensure that relatively low cutoffs can be used, we replace the core electrons with separable norm-conserving GTH (HGH) pseudopotentials [46, 47], which removes the sharp nuclear densities. We use the Gaussian basis sets that are designed for use with these pseudopotentials in solid-state calculations [2].

The one-electron overlap, kinetic energy, and local (L) pseudopotential integrals are
evaluated by numerical integration on the real-space grid according to

\[
S_{\mu\nu}(k) = \int_{\Omega} dr \phi_\mu^*(r) \phi_\nu(k(r)),
\]
(99)

\[
T_{\mu\nu}(k) = -\frac{1}{2} \int_{\Omega} dr \phi_\mu^*(r) \nabla^2 r \phi_\nu(k(r)),
\]
(100)

\[
V_{\mu\nu}^{\text{L}}(k) = \int_{\Omega} dr \phi_\mu^*(r) v^{\text{L}}(r) \phi_\nu(k(r));
\]
(101)

where we note that these integrals and other quantities throughout the paper are defined per unit cell. The total, local contribution from the ion pseudopotentials \(v^{\text{L}}(r)\) is given by,

\[
v^{\text{L}}(r) = \sum_{G \neq 0} \sum_{I} e^{iG \cdot (r - R_I)} v_I^\text{L}(G),
\]
(102)

where \(I\) indexes the ions at position \(R_I\) in the unit cell. The local part of the GTH pseudopotential is finite at \(r = 0\) but decays as \(-Z_I/r\) leading to a divergent \(G = 0\) component, which is treated separately (see below). The nonlocal (NL) pseudopotential part is evaluated on the reciprocal-space grid

\[
V_{\mu\nu}^{\text{NL}}(k) = \Omega \sum_{G, G'} \phi^*_\mu(k(G)) v^{\text{NL}}(k + G, k + G') \phi_\nu(G')
\]
(103)

with

\[
v^{\text{NL}}(k + G, k + G') = \sum_{I} e^{-i(G - G') \cdot R_I} v_I^{\text{NL}}(k + G, k + G')
\]
(104)

and

\[
v_{I}^{\text{NL}}(k + G, k + G') = \frac{1}{\Omega} \int dr \int dr' e^{-i(k + G) \cdot r} \times v_{I}^{\text{NL}}(r, r') e^{i(k + G') \cdot r'}.
\]
(105)

In the above, the ion-specific local and nonlocal parts of the pseudopotential have analytic definitions of the GTH form [46, 47]. In particular the nonlocal part is separable and expressible as a sum of products of functions of \(k + G\) and functions of \(k + G'\).
The Hartree and exchange matrices are evaluated in real space,

\[ J_{\mu\nu}(k) = \int_\Omega dr \phi^*_\mu k(r) v_H(r) \phi_\nu k(r), \quad (106) \]

\[ K_{\mu\nu}(k) = \int_\Omega dr \int dr' \phi^*_\mu k(r) \frac{\rho(r, r')}{|r - r'|} \phi_\nu k(r'), \quad (107) \]

in terms of the Hartree potential

\[ v_H(r) = \frac{4\pi}{\Omega} \sum_{G \neq 0} \frac{\rho(G)}{G^2} e^{iG \cdot r} \quad (108) \]

and the density matrix \( \rho(r, r') \). In general, the density matrix and density \( \rho(r) = \rho(r, r) \) can be obtained from a Brillouin zone sampling

\[ \rho(r, r') = \frac{1}{N_k} \sum_k \sum_{\lambda\sigma} P_{\lambda\sigma}(k) \phi_{\lambda k}(r) \phi^*_{\sigma k}(r'), \quad (109) \]

where \( N_k \) is the number of \( k \)-points sampled in the Brillouin zone. Assuming a closed-shell reference of molecular orbitals, \( \psi_{pk} = \sum_\mu C_{\mu p}(k) \phi_{\mu k} \), the mean-field density matrix is given by

\[ P_{\lambda\sigma}(k) = 2 \sum_i^{\text{occ}} C_{\lambda i}(k) C^*_{\sigma i}(k). \quad (110) \]

This leads to the real-space form of the exchange matrix

\[ K_{\mu\nu}(k) = \frac{1}{N_k} \sum_{k'} \sum_{\lambda\sigma} P_{\lambda\sigma}(k') \int dr \phi^*_\mu k(r) v^X_{\lambda k', \sigma k}(r) \phi_{\lambda k'}(r), \quad (111) \]

where

\[ v^X_{\sigma k', \nu k}(r) = \int dr' \frac{\phi^*_\sigma k'(r') \phi_{\nu k}(r')}{|r - r'|} = \int dr' \frac{\rho_{\sigma k', \nu k}(r')}{|r - r'|} \]

\[ = \frac{4\pi}{\Omega} \sum_G \frac{\rho_{\sigma k', \nu k}(G)}{|k - k' + G|^2} e^{i(k - k' + G) \cdot r}, \quad (112) \]

and the \( G = 0 \) term is excluded when \( k = k' \). Like the individual crystalline AOs, the
AO pair densities have a plane-wave resolution,

\[ \rho_{\sigma k', \nu k}(r) = \phi_{\sigma k'}^*(r) \phi_{\nu k}(r) = \sum_G \rho_{\sigma k', \nu k}(G) e^{i(k-k'+G) \cdot r}, \]  

(113)

\[ \rho_{\sigma k', \nu k}(G) = \frac{1}{\Omega} \int_\Omega \rho_{\sigma k', \nu k}(r) e^{-i(k-k'+G) \cdot r}. \]  

(114)

As shown above, the local part of the pseudopotential and Hartree potential are separately divergent at \( G = 0 \), however their sum is not. For each atom, the sum is given by

\[ \alpha_I = \int d\mathbf{r} \left( v^I_L(\mathbf{r}) + \frac{Z_I e^2}{r} \right), \]  

(115)

where \( \alpha_I \) is a finite, ion-specific parameter of the pseudopotential. This leads to an additional matrix element

\[ V^{L+J}_{\mu \nu}(k) = S_{\mu \nu}(k) \sum_I \alpha_I, \]  

(116)

resulting simply in a uniform shift of the orbital energies and a constant in the total energy. Note that in the absence of pseudopotentials, the divergent contributions from the Hartree and electron-nuclear interaction cancel exactly by charge neutrality.

Once the above integrals are defined, then it is straightforward to carry out a HF calculation using these integrals. The only difference compared to a standard molecular HF calculation is that the integrals and orbitals are complex. The molecular orbitals at each \( k \)-point \( \psi_{\nu k} \) are obtained from the HF equation

\[ \mathbf{F}(k) \mathbf{C}(k) = \varepsilon(k) \mathbf{S}(k) \mathbf{C}(k) \]  

(117)

where

\[ \mathbf{F}(k) = \mathbf{T}(k) + \mathbf{V}^{PP}(k) + \mathbf{J}(k) - \frac{1}{2} \mathbf{K}(k) + \mathbf{V}^{L+J}(k) \]  

(118)

and \( \mathbf{V}^{PP}(k) = \mathbf{V}^L(k) + \mathbf{V}^{NL}(k) \). Assuming \( n_{\text{el}} \) electrons per unit cell, the HF determinant includes the \( N_k n_{\text{el}} \) orbitals with the lowest eigenvalues out of all \( k \)-points. The
total energy is the usual HF one plus the nuclear-nuclear repulsion, which is computed using the Ewald expression

\[
E_{\text{NN}} = \frac{1}{2} \sum_{I \neq J} \frac{Z_I Z_J}{|R_I - R_J - T|} \text{erfc} \left( \eta |R_I - R_J - T| \right) + \frac{1}{2} \sum_{G \neq 0} \left| \sum_{I} Z_I e^{iG \cdot R_I} \right|^2 \frac{e^{-G^2/4\eta^2}}{G^2} - \frac{\eta}{\sqrt{\pi}} \sum_{I} Z_I^2 - \frac{\pi}{2\Omega^2} \left( \sum_{I} Z_I \right)^2,
\]

where the primed summation neglects the self-interaction terms with \( I = J \) when \( T = 0 \) and the range of lattice summations is chosen together with \( \eta \) to facilitate rapid convergence.

For a subsequent correlation treatment, we need to define the molecular orbital integrals. The one-electron integrals can be obtained straightforwardly by changing basis with \( C(k) \). Two-electron integrals are defined (again, per unit cell) as

\[
\langle pk_p, rk_r | qk_q, sk_s \rangle = \Omega \sum_{G} \left[ \beta_{pk_p, qk_q}(G) \right] \times v(k_q - k_p + G)r_{rk_r, sk_s}(G_{prqs} - G).
\]

In the presence of translational invariance, these two-electron integrals must conserve crystal momentum, i.e. \( k_p + k_r - k_q - k_s = G \), where \( G \) is a reciprocal lattice vector. We evaluate these integrals by representing the molecular orbital pair density in terms of plane-waves, leading to

\[
\langle pk_p, qk_q | rk_r, sk_s \rangle = \Omega^2 \sum_{G} \left[ \beta_{pk_p, qk_q}(G) \right] \times v(k_q - k_p + G)r_{rk_r, sk_s}(G_{prqs} - G).
\]
correlation techniques, it is convenient to express the Hamiltonian in normal-ordered form

\[ H = \langle \Phi | H | \Phi \rangle + \sum_{pq} \sum_k F_{pq}(k) \left\{ E_{qk}^p \right\} + \frac{1}{2} \sum_{pqrs} \sum_{k} \left( \rho_{pq} \dot{\rho}_{sk} \right) \left\{ E_{qk}^p E_{sk}^r \right\} \]

where \( E_{qk}^p = \sum_{\sigma} a_{pk\sigma}^\dagger a_{qk\sigma} \) is a spin-summed excitation operator, \{\ldots\} denotes normal ordering with respect to the reference \( \Phi \) (the Fermi vacuum), and as usual the primed summation enforces crystal momentum conservation. Naturally, in the case of canonical HF, the Fock matrix is diagonal, \( F_{pq}(k) = \varepsilon_p(k) \delta_{pq} \).

### 6.3 Periodic EOM coupled-cluster

Given a closed-shell reference determinant \( |\Phi\rangle \), the ground-state CCSD wavefunction is \( |\Psi_0\rangle = e^T |\Phi\rangle \) where \( T = T_1 + T_2 \), and \( T_1 \) and \( T_2 \) are single- and double-excitation operators

\[
T_1 = \sum_{ai} \sum_{k_0k_i} i_{ik_0}^{a_k a_i} E_{ik_i}^{a_k a_i}, \\
T_2 = \frac{1}{2} \sum_{abij} \sum_{k_0k_j k_i k_l} i_{ik_0j}^{a_k a_l} E_{ik_i}^{a_k a_l} E_{jk_j}^{a_l a_i} E_{ik_l}^{a_l a_i} E_{jk_i}^{a_i a_i}\]

In this work, we use a HF reference determinant, although the equations below and in the Appendix apply to any determinant. In Eqs. (123) and throughout, the indices \( i, j, k, l \) denote occupied orbitals and \( a, b, c, d \) denote virtual orbitals. Because the crystal Hamiltonian has translational symmetry, the excitation operators must conserve crystal momentum, i.e. \( \sum_a k_a - \sum_i k_i = G \) where \( k_a \) and \( k_i \) are the crystal momenta of particle and hole orbitals and \( G \) is a reciprocal lattice vector. This requirement is indicated by the primed summation in Eq. (123) and we emphasize that each primed summation indicates that one of the listed momenta is fixed and need not be summed. Introducing the non-Hermitian CC effective Hamiltonian \( \tilde{H} = e^{-T} H e^T \), the CCSD ground-state
energy and excitation amplitudes are determined by

\[ E_0 = \langle \Phi_0 | \hat{H} | \Phi_0 \rangle, \quad (124) \]

\[ 0 = \langle \Phi_{ik_1,\alpha}^a | \hat{H} | \Phi_0 \rangle, \quad (125a) \]

\[ 0 = \langle \Phi_{ik_1,\alpha, jk_2,\beta}^{a,b} | \hat{H} | \Phi_0 \rangle, \quad (125b) \]

where \( \Phi_{ik_1,\alpha}^a \) and \( \Phi_{ik_1,\alpha, jk_2,\beta}^{a,b} \) are Slater determinants with one and two electron-hole pairs. The explicit forms of the energy and amplitude equations with translational symmetry are given in Ref. [12]. Importantly, the computational cost of CCSD for periodic systems scales as \( n_{\text{occ}}^2 n_{\text{vir}}^4 N_k^4 \), where \( n_{\text{occ}} \) and \( n_{\text{vir}} \) denote the number of occupied and virtual orbitals per unit cell; this is a factor of \( N_k^2 \) less than the equivalent calculation that neglects momentum conservation.

Coupled-cluster excited states and energies are determined through the equation-of-motion (EOM) formalism [32, 67, 68], which amounts to diagonalizing the effective Hamiltonian \( \hat{H} \) in an appropriate space of excitations. For a CCSD ground state, we calculate ionization potentials (IPs) via diagonalization in the space of 1-hole (1h) and 2-hole, 1-particle (2h1p) states, and we calculate electron affinities (EAs) via diagonalization in the space of 1-particle (1p) and 2-particle, 1-hole (2p1h) states:

\[ |\Psi_{n,k}^{N\pm1}\rangle = R_1^\pm (n,k) |\Psi_0\rangle = [R_1^+ (n,k) + R_1^- (n,k)] |\Psi_0\rangle \]

where \( R_1^- (R_1^+) \) creates 1h (1p) excitations and \( R_2^+ \) (\( R_2^- \)) creates 2h1p (2p1h) excitations in the IP-EOM (EA-EOM) framework. Explicitly,

\[ R_1^-(n,k) = \sum_i r_{i,k} a_{i,k_\alpha} + \sum_{aij} r_{ik_1,jk_1}^{a,k} E_{jk_1}^{a,k} a_{i,k_\alpha} \quad (126) \]

\[ R_1^+(n,k) = \sum_a r_{a,k} a_{a,k_\alpha}^\dagger + \sum_{abij} r_{ik_1,jk_1}^{a,b} E_{jk_1}^{a,b} a_{a,k_\alpha}^\dagger \quad (127) \]

where again the primed summations are restricted to enforce momentum conservation.

The \( r \)-amplitudes satisfy the usual IP/EA EOM equations, modified to include crystal momenta; explicit expressions are given in the Appendix. Again we emphasize that each primed summation leads to one crystal momentum which can be fixed by momentum...
conservation; for example, in the final term of the $R_2$ amplitude equations (133) and (135), only two of the four crystal momenta need to be explicitly summed. Note that although ground-state CCSD scales as $N^6$, the subsequent EOM-CCSD calculations have a reduced scaling of $N^5$. In particular, in the presence of periodicity, the EOM-CCSD scaling for all charged excitations at a given $k$-point is $n_{\text{occ}}^3 n_{\text{vir}}^2 N_k^3$ for IP-EOM-CCSD and $n_{\text{occ}}^4 n_{\text{vir}}^3 N_k^3$ for EA-EOM-CCSD (if the charged excitations at all sampled $k$-points are desired, then $N_k$ such calculations can be performed independently). This specific scaling with the number of occupied and virtual orbitals usually makes the calculation of conduction (virtual) bands significantly more expensive than that of valence (occupied) bands.

In this work we focus on the quasiparticle excitations. These are the many-body states that have the largest overlap with the mean-field single-particle excitations, i.e. the largest $|r_{i,k}|^2$ and $|r_{a,k}|^2$ (a proxy for the EOM Green’s function pole strength). This observation leads to an efficient targeted Davidson diagonalization procedure based on the wavefunction character, rather than the energy [34, 123]. At each $k$-point sampled in the Brillouin zone, we typically target the three lowest-lying IP and EA excitations with such single-particle character. Away from the band edge, the quasiparticle picture breaks down: the excitations develop an effective linewidth and nontrivial satellite structure through a growing contribution of multiple determinants, which also leads to slower convergence of the Davidson procedure. In this regime, it is more efficient to directly construct the momentum- and frequency-dependent, one-particle Green’s function as done in our previous work on the uniform electron gas [21].

### 6.4 Convergence and finite-size effects

It is worth briefly discussing the requirements for converging the correlated electronic structure of a periodic system. As in a molecular electronic structure calculation, we must converge with respect to the single particle basis set, i.e. the set of Gaussian functions $\tilde{\phi}_\mu(r)$ in Eq. (97), as well as with respect to the correlation level. Convergence in these two regards can be expected to be similar to that in a molecular system.
However, in addition, we must converge with respect to the Brillouin zone sampling, i.e. the set of crystal momentum vectors $\mathbf{k}$ appearing in Eq. (97). (This may formally be thought of as part of the basis set convergence for the infinite system). There is substantial experience with this convergence behavior at the mean-field level (HF and DFT), but much less so at the correlated level [112–114, 124]. However, it is important to note that convergence with respect to Brillouin zone sampling can be quite slow.

In general, if all quantities are smooth in the Brillouin zone, then approximating integrals by finite Brillouin zone sampling has exponentially small error in the grid spacing, i.e. $\exp(-aN_k^{-1/3})$. However, omitting the $G = 0$ contribution in the definition of Coulomb-based integrals is an additional source of finite-size error. For example, the exact HF exchange energy per unit cell is given by

$$
E_X = -\sum_{ij} \int_{BZ} \frac{\Omega d\mathbf{k}_i}{(2\pi)^3} \int_{BZ} \frac{\Omega d\mathbf{k}_j}{(2\pi)^3} (i\mathbf{k}_i j\mathbf{k}_j j\mathbf{k}_j i\mathbf{k}_i). 
$$

In this limit of infinite $k$-point sampling, the volume element associated with Brillouin zone integration cancels the divergent Coulomb term, i.e. it is an integrable divergence. However, using the molecular orbital integrals from Eq. (121) and a finite Brillouin zone sampling leads to the approximate exchange energy

$$
E_X = -\frac{\Omega}{N_k^2} \sum_{ij} \left| \rho_{ik,i} j\mathbf{k}_j (G = 0) \right|^2 \frac{4\pi}{|G - q_{ij}|^2},
$$

where the primed summation excludes $G = 0$ when $q_{ij} = \mathbf{k}_i - \mathbf{k}_j = 0$. This neglected term represents an $O(1)$ integrand coming from the nonzero charge density $\rho_{ik,ik}(G = 0)$, leading to an integration error of $O(N_k^{-1/3})$. In these cases related to HF exchange, various corrections exist that aim to accelerate the Brillouin zone convergence, including auxiliary function techniques [125] and real-space truncation of the Coulomb interaction [126, 127]. In this work, all correlated calculations and their underlying HF calculations are done with no such corrections. However, when total energies are required, we separately converge the HF energy using an exchange interaction with a spherically truncated Coulomb potential; these calculations are performed
with isotropic $k$-point meshes and the results are extrapolated using a finite-size scaling of the form $\exp(-aN_k^{1/3})$ [126, 127].

Finite-size errors in the correlated theories can be analyzed in a similar manner. With $k$-point sampling, the MP2 energy per unit cell is approximated by

$$E_{\text{MP2}} = \frac{\Omega}{N_k^3} \sum_{k_a,k_b,k_i,k_j} \sum_{abij} t_{ak_{i}bk_{j}}^{ak_{a}bk_{b}}$$

$$\times \left[ 2 \sum_{G} \rho_{ki,ak_{a}}(G) \frac{4\pi}{|G-q_{ia}|^2} \rho_{kj,bk_{b}}(G_{ij} - G) - \sum_{G} \rho_{ki,bk_{b}}(G) \frac{4\pi}{|G-q_{ib}|^2} \rho_{kj,ak_{a}}(G_{ij} - G) \right]$$

(130)

where again the term with $G = 0$ is neglected when $q_{ia} = k_i - k_a = 0$ (first term) or $q_{ib} = k_i - k_b = 0$ (second term). Unlike in the case of exchange, these neglected terms vanish at the origin $q = 0$ because the orthogonality of orbitals guarantees that $\rho_{ik,ak}(G = 0) = 0$; instead, these terms are associated with an integrand of $O(q^2)$, leading to an integrated error of $O(N_k^{-1})$. If the $t$-amplitudes are correct, then this is the only integration error. However, the amplitudes

$$t_{ak_{i}bk_{j}}^{ak_{a}bk_{b}} = \frac{(ik_{i}ak_{a})^{*}(jk_{j}bk_{b})^{*}}{\varepsilon_{ik_{i}} + \varepsilon_{jk_{j}} - \varepsilon_{ak_{a}} - \varepsilon_{bk_{b}}}$$

(131)

inherit the error of the HF orbital energies, which can be analyzed analogously. Orthogonality guarantees that unoccupied energies have the same favorable $N_k^{-1}$ error; however, just as for the exchange energy, the occupied orbitals exhibit an $N_k^{-1/3}$ error [128]. Without any corrections to the orbital energies, this pollutes the $t$-amplitudes and thus dominates the error in the correlation energies (see below). The inclusion of all four-index integrals in CCSD is similarly responsible for a $N_k^{-1/3}$ error.

### 6.5 Ground-state results

We first consider the ground-state properties of diamond and silicon, to establish the convergence properties of the ground state CC before proceeding to our excited state studies in the next section. Diamond and silicon share the same crystal structure, with
two atoms per (primitive) unit cell. Except where otherwise noted, all calculations are performed with the $T = 300$ K experimental lattice constants $a = 3.567$ Å and $a = 5.431$ Å, for diamond and silicon respectively.

We use GTH pseudopotentials [46, 47], explicitly treating four valence electrons per atom (eight per unit cell), and matching single-particle basis sets, obtained from the CP2K software package [1, 2, 129]. In the present work, we use pseudopotentials and basis sets that were originally optimized for use in DFT calculations with the local density approximation (LDA); future work will consider pseudopotentials and basis sets that are optimized for HF calculations [115] and we will present a comparison with all-electron calculations. The DZV, DZVP, and TZVP basis sets have 8, 13, and 17 orbitals per atom (twice as many per cell). The real-space grid spacing used for the integrals was about 0.17 Å, corresponding to a kinetic energy cutoff of approximately 400 Ry (note that the same grid is used to resolve the orbitals, Eq. (98), and the pair densities, Eq. (113)). All calculations are converged, with respect to this grid, to an accuracy better than $10^{-4}$ au per atom. In all ground-state calculations, the Brillouin zone was sampled from a uniform Γ-centered grid. All calculations were performed with the PySCF software package [10].

As described above, total convergence in periodic solids must be achieved with respect to the single-particle basis and the sampling of the Brillouin zone (finite-size effects). In Fig. 3 we show the convergence of the ground-state correlation energy of diamond, at the MP2 and CCSD levels of theory, as a function of the size of the single-particle basis and the number of $k$-points sampled in the Brillouin zone. As expected from molecular calculations with comparable basis sets, the total correlation energies are not converged with respect to the single-particle basis; the correlation energy difference between results obtained with the DZVP and TZVP basis sets is 0.4–0.5 eV per unit cell (half that per atom), to be compared to the HF energy differences, which are less than 0.1 eV (not shown). In this work, we do not pursue a complete basis set limit extrapolation based only on double-zeta and triple-zeta basis sets. It can be expected, however, that this basis set error is local in character and can be corrected
<table>
<thead>
<tr>
<th></th>
<th>DZV</th>
<th>DZVP</th>
<th>TZVP</th>
</tr>
</thead>
<tbody>
<tr>
<td>HF</td>
<td>Extrap, exp((-aN_k^{1/3}))</td>
<td>-301.05</td>
<td>-301.86</td>
</tr>
<tr>
<td>MP2</td>
<td>3 × 3 × 3</td>
<td>-5.91</td>
<td>-8.27</td>
</tr>
<tr>
<td></td>
<td>4 × 4 × 4</td>
<td>-5.96</td>
<td>-8.24</td>
</tr>
<tr>
<td></td>
<td>Extrap, (N_k^{-1/3})</td>
<td>-6.10</td>
<td>-8.16</td>
</tr>
<tr>
<td>CCSD</td>
<td>3 × 3 × 3</td>
<td>-5.01</td>
<td>-7.00</td>
</tr>
<tr>
<td></td>
<td>4 × 4 × 4</td>
<td>-5.15</td>
<td>-7.12</td>
</tr>
<tr>
<td></td>
<td>Extrap, (N_k^{-1/3})</td>
<td>-5.58</td>
<td>-7.50</td>
</tr>
</tbody>
</table>

Table 1: Hartree-Fock total energy per unit cell and MP2/CCSD correlation energy per unit cell of diamond in eV.

using standard explicit \(r_{12}\) correlation techniques [130]. Furthermore, we observe that the basis set corrections to the correlation energy are similar for MP2 and CCSD (as in molecules), and we could thus similarly use complete basis set limit MP2 calculations to correct the CCSD results.

With regards to finite-size effects, we observe that in small basis sets the MP2 correlation energy converges more quickly than the CCSD energy with the number of \(k\)-points. However, for the largest (TZVP) basis set, the energy convergence is similar. In Tab. 1, we give the HF and correlation energies obtained with a 3 × 3 × 3 and 4 × 4 × 4 sampling of the Brillouin zone; the largest CC calculation presented here, using a 4 × 4 × 4 sampling of the Brillouin zone and the TZVP basis, constitutes a canonical CCSD calculation of 256 electrons in 2,176 orbitals, demonstrating the savings provided by incorporating periodic symmetry.

In Tab. 1, we also report the thermodynamic limit based on an extrapolation of the form \(N_k^{-1/3}\), as discussed in Sec. 6.4. This extrapolation is graphically demonstrated in Fig. 4, which shows the cohesive energy (the energy-per-atom difference between the crystal and isolated atom) of diamond and silicon as a function of the number of \(k\)-points sampled using the TZVP basis. Specifically, the fitting is restricted to the two largest isotropic \(k\)-point meshes, \(n \times n \times n\) with \(n = 3, 4\). Atomic energies of open-shell
a (Å) & B (GPa) & ΔE (eV/atom)
\hline
C & & \\
HF & 3.527 & (3.552) & 507 & (495) & 5.36 & (5.28) \\
MP2 & 3.545 & (3.553) & 436 & (450) & 7.91 & (7.97,8.039) \\
CCSD & 3.539 & — & 463 & — & 7.04 & (7.295) \\
Experiment & 3.553 & — & 455 & — & — & 7.55 \\
Si & & \\
HF & 5.435 & (5.512) & 107 & (103) & 3.03 & (2.97) \\
MP2 & 5.347 & (5.415) & 101 & (100) & 4.96 & (5.05) \\
CCSD & 5.393 & — & 103 & — & 4.15 & — \\
Experiment & 5.421 & — & 101 & — & — & 4.68 \\
\hline

Table 2: Lattice constant \( a \), bulk modulus \( B \), and cohesive energy \( \Delta E \) of diamond and silicon using the TZVP single-particle basis. Lattice constant and bulk modulus were obtained from fits to calculations with a \( 3 \times 3 \times 3 \) \( k \)-point mesh. Crystal energy contributions to the cohesive energy are calculated at the \( T = 300 \) K lattice constant and extrapolated to \( N_k \rightarrow \infty \). Values given in parentheses are those calculated with the plane-wave PAW approach as reported in Refs. [8, 113]. All experimental values have been corrected for zero-point vibrational effects [131].

C and Si were calculated with unrestricted MP2 and CCSD in the field of crystalline basis functions (to account for basis set superposition error) and the HF contribution to the crystal energy is separately converged with respect to the number of \( k \)-points, in the manner described in Sec. 6.4. Predicted cohesive energies are given in Tab. 2.

Our cohesive energies for diamond, 7.91 eV/atom (MP2) and 7.04 eV/atom (CCSD), are in reasonable agreement with those of previous periodic MP2 calculations, 7.97 eV/atom [113] and 8.039 eV/atom [8], and of a previous CCSD calculation, 7.295 eV/atom [8]. Similarly, our MP2 cohesive energy for silicon, 4.96 eV/atom, is in good agreement with the previously reported MP2 value, 5.05 eV/atom [113]. The discrepancies in our data of order 0.1–0.2 eV/atom likely originate from a combination of the finite basis set and treatment of core electrons. The calculations reported in Refs. [8, 113] employ a plane-wave basis set and the projector-augmented wave (PAW) method, which differs from the norm-conserving pseudopotentials used here. Furthermore, we always find that the absolute value of the correlation energy given by MP2 is larger than that given by CCSD, which is the opposite of that observed in Ref. [8]; this difference may also be attributable to the difference between the PAW and pseudopotential
approximations.

Using a $3 \times 3 \times 3$ $k$-point mesh and the TZVP basis, we have further calculated the equation of state of diamond and silicon, shown in Fig. 5. We performed at least seven ground-state calculations with unit cell volumes varying by about $\pm 10\%$ from the equilibrium value and fit the results to a third-order Birch-Murnaghan form, leading to predicted (zero-temperature) lattice constants and bulk moduli, given in Tab. 2. Again, the HF and MP2 structural predictions are in good agreement with previous values [113], and the observed discrepancy at the HF level provides some measure of the error incurred due to the use of pseudopotentials. Ultimately, the overall good agreement between our results and those of Refs. [8, 113] suggests that the correlation energy differences required for structural properties and cohesive energies are reasonably converged in both implementations, with respect to both the basis and $k$-point sampling. At least for the two semiconductors studied here, MP2 and CCSD provide a similar level of accuracy when compared to experimental values determined by ground-state energetics.

6.6 Excited-state results

We now turn to the excited-state and spectral calculations using the EOM-CCSD formalism. As discussed in Sec. 6.3, we compute the quasiparticle excitations in the ionization potential (IP) and electron affinity (EA) charged sectors. To obtain continuous bands in conventional band structure calculations with HF or DFT, an initial self-consistent calculation is performed with a fixed (and potentially coarse) $k$-point grid, which provides an approximate representation of the density matrix or density; see Eq. (109). From this latter object, the one-body Fock-like matrix can be constructed at arbitrary $k$-points and diagonalized, leading to non-self-consistent eigenvalues. To construct a continuous band-structure in EOM-CCSD, we here perform a large series of (independent) calculations with a $k$-point mesh that is shifted to include the desired $k$-point. Although this means that certain calculations along the band path will be performed with a sub-optimal mesh, the results will converge properly in the limit
of dense $k$-point sampling. We only present results here for a band structure path through the Brillouin zone, but the same technique could be applied to calculate the full quasiparticle density of states.

The band structures of diamond and silicon are shown in Figs. 6 and 7 using DFT with the PBE functional [132], HF, and EOM-CCSD using the DZVP basis and a $3 \times 3 \times 3$ $k$-point mesh. Qualitatively, the EOM-CCSD result reproduces the expected behavior: it predicts a band gap that is slightly larger than that of PBE and significantly smaller than that of HF. Similarly, the EOM-CCSD band widths are intermediate between the two mean-field results.

To study the convergence of the excited-state structure, we present in Fig. 8 the calculated value of the indirect (minimum) band gap in diamond and silicon, as a function of the basis set and number of $k$-points sampled. Since we do not calculate the full band path due to its high cost, the indirect band gap is evaluated from two independent calculations: an IP-EOM-CCSD calculation with a $\Gamma$-centered mesh and an EA-EOM-CCSD calculation with a $k$-point mesh centered at $\sim 85\%$ of the $\Gamma - X$ line (near the zone boundary). We see that larger single-particle basis sets favor a smaller bandgap and finer $k$-point meshes favor a larger band gap. The band gap difference between calculations using the DZVP and TZVP basis sets is about 0.1 eV, and relatively independent of the number of $k$-points sampled.

The band gap convergence with the number of $k$-points is slow and oscillatory (when including anisotropic $k$-point meshes), a behavior inherited from the underlying HF band structure (not shown). In particular, we retain the slow convergence arising from the $G = 0$, $q \rightarrow 0$ contributions to the energy discussed in Sec. 6.2, which appears to lead to an $N_k^{-1/3}$ scaling of the error in the bandgap. Similarly to the ground-state energy however, the finite-size effects are much smaller in the correlated band gap than in the HF band gap: in diamond, the HF band gap changes by about 7 eV during convergence, while the EOM-CCSD band gap changes by less than 1 eV.

Like in our treatment of the ground-state and cohesive energy, we have applied a $N_k^{-1/3}$ finite-size scaling to the EOM-CCSD band gaps, restricted to the two largest
isotropic $k$-point meshes. Within the approximations of our approach (GTH pseudopotentials, finite basis, and EOM-CCSD correlation) we predict band gaps of 5.37 eV and 1.19 eV, for diamond and silicon respectively; the experimental $T = 0$ band gaps are 5.45–5.50 eV [133, 134] and 1.17 eV [134, 135], which are corrected for excitonic and finite-temperature effects, but not for the more challenging electron-phonon (vibrational) effects. For diamond in particular, a zero-point band gap renormalization of 0.4–0.6 eV has been estimated [136–138], which would suggest an experimental electronic band gap of 5.8 eV or more. For comparison, tightly converged all-electron $G_0W_0$ calculations predict band gaps of 5.61–5.62 eV and 1.10–1.11 eV [139, 140], which are quite close to $G_0W_0$ calculations with norm-conserving pseudopotentials [141], but 0.1–0.2 eV smaller than all-electron self-consistent $GW$ calculations [142] and 0.2–0.3 eV smaller than vertex-corrected calculations [105]. Interestingly, although MP2 predicts ground-state structural properties that are on par with CCSD, it massively underestimates the band gaps of these two materials, giving 1.9 eV for diamond and $-1.2$ eV for silicon [113], which was attributed to the large polarizability of these materials.

6.7 Discussion and conclusion

In this work, we have presented the first-ever results from canonical Gaussian-based coupled-cluster theory for three-dimensional solids. We discussed concomitant finite-size errors in the mean-field and correlated calculations. Based on ground-state coupled-cluster calculations, we presented the lattice constant, bulk modulus, and cohesive energies of diamond and silicon; these results are in good agreement with limited existing data based on the plane-wave PAW approach [8, 113]. We further presented the first results of excited-state (equation-of-motion) coupled-cluster theory for the band structure and band gap of the same two semiconductors. Our predicted minimum band gaps, 5.37 eV for diamond and 1.19 eV for silicon, are in good agreement with experimental values.

A great deal of work remains to be done to make coupled-cluster calculations on weakly-correlated solids routine. In future work, we will carefully compare the results of
pseudopotential and all-electron calculations, which represents an unquantified source of error. Parallel efforts are aimed at reducing the cost and increasing our understanding of these calculations, for example through local correlation approaches, explicitly correlated formulations, perturbative corrections due to triple excitations, and applications to a wider range of insulating and metallic materials.

6.8 IP/EA-EOM-CCSD equations for periodic systems

Assuming a closed-shell reference, the spatial-orbital IP-EOM-CCSD amplitude equations are given by

\[
(HR^-)_{ik_i} = \sum_{k} \sum_{k_b} U_{kk_ik_i} r_{kk_b} + \sum_{ld} \sum_{k_b} U_{kk_i} \alpha_{ik_b} \left( 2r_{ik_b} - r_{ik_b} \right)
\]

\[
- \sum_{kld} \sum_{kb} \left( 2W_{kk_ik_i} W_{kk_ik_i} - W_{kk_ik_i} \right) r_{kb}
\]

(132)

\[
(HR^-)_{ik_b} = \sum_{k} \sum_{kb} U_{kk_ik_b} r_{kb} + \sum_{l} \sum_{kb} U_{kk_i} \alpha_{lk_b} \left( 2r_{lk_b} - r_{lk_b} \right)
\]

\[
- \sum_{lk} \sum_{lk_b} \left( 2W_{lk_ik_b} W_{lk_ik_b} - W_{lk_ik_b} \right) r_{lk_b}
\]

\[
+ \sum_{kl} \sum_{kb} W_{lk_ik_b} \alpha_{lk_b} \left( 2W_{lk_ik_b} - W_{lk_ik_b} \right) r_{lk_b}
\]

\[
- \sum_{l} \sum_{lk} \sum_{lk_b} \left( 2W_{lk_ik_b} W_{lk_ik_b} - W_{lk_ik_b} \right) r_{lk_b}
\]

and the EA-EOM-CCSD equations are given by

\[
(HR^+)_{ak_a} = \sum_{c} \sum_{kc} U_{ak_a} \alpha_{kc} c_{ik} + \sum_{ld} \sum_{k_b} U_{kk_i} \alpha_{ik_b} \left( 2r_{ik_b} - r_{ik_b} \right)
\]

\[
+ \sum_{l} \sum_{kb} \left( 2W_{lk_ik_b} W_{lk_ik_b} - W_{lk_ik_b} \right) r_{lk_b}
\]

(134)
As described in the text, primed summations indicate that one of the listed momenta is fixed by momentum conservation. The intermediates $U$ and $W$ are the usual ones arising in coupled-cluster theory from the effective Hamiltonian $\hat{H}$ [143], and their efficient calculation also accounts for momentum conservation.
Figure 3: Correlation energy per unit cell of diamond calculated with MP2 (left, open symbols) and CCSD (right, filled symbols). Dashed lines indicate the $N_k \to \infty$ extrapolated values assuming a finite-size scaling of the form $N_k^{-1/3}$.

Figure 4: Cohesive energy of diamond (left) and silicon (right), calculated with the TZVP basis. For each $k$-point mesh, MP2 and CCSD correlation energies are added to a separately converged HF crystal energy. The correlation component of the cohesive energy is extrapolated assuming a finite-size error that scales like $N_k^{-1/3}$. Zero-point corrected experimental results [131] are shown as a thin grey line. Predicted and experimental values are given in Tab. 2.
Figure 5: Equation of state of diamond (left) and silicon (right) with the TZVP basis and a $3 \times 3 \times 3$ sampling of the Brillouin zone.

Figure 6: Band structure of diamond calculated with DFT (PBE), HF, and EOM-CCSD, using the DZVP single-particle basis and a $3 \times 3 \times 3$ $k$-point mesh.
**Figure 7:** The same as in Fig. 6, but for silicon.

**Figure 8:** Indirect band gap of diamond (left) and silicon (right) calculated with EOM-CCSD. Dashed lines indicate extrapolation of the form $N_k^{-1/3}$ based on the isotropic $3 \times 3 \times 3$ and $4 \times 4 \times 4$ sampling meshes; grey horizontal lines indicate experimental values. For the TZVP data, the extrapolation predicts for diamond a band gap of 5.37 eV compared to the experimental value of 5.45–5.50 eV [133, 134], and for silicon a band gap of 1.19 eV, compared to the experimental value of 1.17 eV [134, 135].
7 PySCF: The Python-based Simulations of Chemistry Framework

This next section brings together the work on periodic systems built on top of the computational framework made by users in Garnet Chan’s group at Princeton University and Caltech. Our hope is that the framework we provide here in the easy-to-use Pythonic language will excel development in not only in periodic systems but in all areas of computational chemistry. The capabilities of the PySCF program will be explored in both the molecular and periodic settings vii.

7.1 Introduction

The Python programming language is playing an increasingly important role in scientific computing. As a high level language, Python supports rapid development practices and easy program maintenance. While programming productivity is hard to measure, it is commonly thought that it is more efficient to prototype new ideas in Python, rather than in traditional low-level compiled languages such as Fortran or C/C++. Further, through the use of the many high-quality numerical libraries available in Python –

viiThis work is based primarily off our previously published work [10].
such as NumPy\cite{144}, SciPy\cite{145}, and MPI4Py\cite{146} – Python programs can perform at competitive levels with optimized Fortran and C/C++ programs, including on large-scale computing architectures.

There have been several efforts in the past to incorporate Python into electronic structure programs. Python has been widely adopted in a scripting role: the Psi4\cite{147} quantum chemistry package uses a custom “Psithon” dialect to drive the underlying C++ implementation, while general simulation environments such as ASE\cite{148} and PyMatGen\cite{149} provide Python frontends to multiple quantum chemistry and electronic structure packages, to organize complex workflows\cite{150}. Python has also proved popular for implementing symbolic second-quantized algebra and code generation tools, such as the Tensor Contraction Engine\cite{151} and the SecondQuantizationAlgebra library\cite{152, 153}.

In the above cases, Python has been used as a supporting language, with the underlying quantum chemistry algorithms implemented in a compiled language. However, Python has also seen some use as a primary implementation language for electronic structure methods. PyQuante\cite{154} was an early attempt to implement a Gaussian-based quantum chemistry code in Python, although it did not achieve speed or functionality competitive with typical packages. Another early effort was the GPAW\cite{155} code, which implements the projector augmented wave formalism for density functional theory, and which is still under active development in multiple groups. Nonetheless, it is probably fair to say that using Python as an implementation language, rather than a supporting language, remains the exception rather than the rule in modern quantum chemistry and electronic structure software efforts.

With the aim of developing a new highly functional, high-performance computing toolbox for the quantum chemistry of molecules and materials implemented primarily in the Python language, we started the open-source project “Python-based Simulations of Chemistry Framework” (PySCF) in 2014. The program was initially ported from our quantum chemistry density matrix embedding theory (DMET) project\cite{156} and contained only the Gaussian integral interface, a basic Hartree-Fock solver, and a
few post-Hartree-Fock components required by DMET. In the next 18 months, multi-
configurational self-consistent-field (MCSCF), density functional theory and coupled
cluster theory, as well as relevant modules for molecular properties, were added into the
package. In 2015, we released the first stable version, PySCF 1.0, wherein we codified
our primary goals for further code development: to produce a package that emphasizes
simplicity, generality, and efficiency, in that order. As a result of this choice, most
functions in PySCF are written purely in Python, with a very limited amount of C
code only for the most time-critical parts. The various features and APIs are designed
and implemented in the simplest and most straightforward manner, so that users can
easily modify the source code to meet their own scientific needs and workflow. However,
although we have favored algorithm accessibility and extensibility over performance,
we have found that the efficient use of numerical Python libraries allows PySCF to
perform at least as fast as the best existing quantum chemistry implementations. In
this article, we highlight the current capabilities and design philosophy of the PySCF
package.

7.2 Capabilities

Molecular electronic structure methods are typically the main focus of quantum chem-
istry packages. We have put significant effort towards the production of a stable,
feature-rich and efficient molecular simulation environment in PySCF. In addition to
molecular quantum chemistry methods, PySCF also provides a wide range of quantum
chemistry methods for extended systems with periodic boundary conditions (PBC). Ta-
ble 3 lists the main electronic structure methods available in the PySCF package. More
detailed descriptions are presented in subsection 7.2.1 - subsection 7.2.7. Although not
listed in the table, many auxiliary tools for method development are also part of the
package. They are briefly documented in subsection 7.2.8 - subsection 7.2.13.

7.2.1 Self-consistent field methods

Self-consistent field (SCF) methods are the starting point for most electronic structure
calculations. In PySCF, the SCF module includes implementations of Hartree-Fock
(HF) and density functional theory (DFT) for restricted, unrestricted, closed-shell and open-shell Slater determinant references. A wide range of predefined exchange-correlation (XC) functionals are supported through a general interface to the Libxc[157] and XCFUN[158] functional libraries. Using the interface, as shown in Figure 9, one can easily customize the XC functionals in DFT calculations. PySCF uses the LIBCINT[159] Gaussian integral library, written by one of us (QS) as its integral engine. In its current implementation, the SCF program can handle over 5000 basis functions on a single symmetric multiprocessing (SMP) node without any approximations to the integrals. To obtain rapid convergence in the SCF iterations, we have also developed a second order co-iterative augmented Hessian (CIAH) algorithm for orbital optimization [160]. Using the direct SCF technique with the CIAH algorithm, we are able to converge a Hartree-Fock calculation for the open-shell molecule Fe(II)-porphine (2997 AOs) on a 16-core node in one day.

7.2.2 Post SCF methods

Single-reference correlation methods can be used on top of the HF or DFT references, including Møller-Plesset second-order perturbation theory (MP2), configuration interaction, and coupled cluster theory.

Canonical single-reference coupled cluster theory has been implemented with single and double excitations (CCSD)[12, 161–163] and with perturbative triples [CCSD(T)][164, 165]. The associated derivative routines include CCSD and CCSD(T) density matrices, CCSD and CCSD(T) analytic gradients[166, 167], and equation-of-motion CCSD for the ionization potentials, electron affinities, and excitation energies (EOM-IP/EA/EE-CCSD)[35, 83, 168–170]. The package contains two complementary implementations of each of these methods. The first is a set of straightforward spin-orbital and spatial-orbital implementations, which are optimized for readability and written in pure Python using the syntax of the NumPy einsum function (which can use either the default NumPy implementation or a custom gemm-based version) for tensor contraction. These implementations are easy for the user to modify. A second spatial-orbital implementa-
tion has been intensively optimized to minimize dataflow and uses asynchronous I/O and a threaded gemm function for efficient tensor contractions. For a system of 25 occupied orbitals and 1500 virtual orbitals (H$_{50}$ with cc-pVQZ basis), the latter CCSD implementation takes less than 3 hours to finish one iteration using 28 CPU cores.

The configuration interaction code implements two solvers: a solver for configuration interaction with single and double excitations (CISD), and a determinant-based full configuration interaction (FCI) solver[171] for fermion, boson or coupled fermion-boson Hamiltonians. The CISD solver has a similar program layout to the CCSD solver. The FCI solver additionally implements the spin-squared operator, second quantized creation and annihilation operators (from which arbitrary second quantized algebra can be implemented), functions to evaluate the density matrices and transition density matrices (up to fourth order), as well as a function to evaluate the overlap of two FCI wavefunctions in different orbital bases. The FCI solver is intensively optimized for multi-threaded performance. It can perform one matrix-vector operation for 16 electrons and 16 orbitals using 16 CPU cores in 30 seconds.

### 7.2.3 Multireference methods

For multireference problems, the PySCF package provides the complete active space self consistent field (CASSCF) method[172, 173] and N-electron valence perturbation theory (NEVPT2)[174, 175]. When the size of the active space exceeds the capabilities of the conventional FCI solver, one can switch to external variational solvers such as a density matrix renormalization group (DMRG) program[176–178] or a full configuration interaction quantum Monte Carlo (FCIQMC) program[179, 180]. Incorporating external solvers into the CASSCF optimizer widens the range of possible applications, while raising new challenges for an efficient CASSCF algorithm. One challenge is the communication between the external solver and the orbital optimization driver; communication must be limited to quantities that are easy to obtain from the external solver. A second challenge is the cost of handling quantities associated with the active space; for example, as the active space becomes large, the memory required to hold
integrals involving active labels can easily exceed available memory. Finally, any approximations introduced in the context of the above two challenges should not interfere with the quality of convergence of the CASSCF optimizer.

To address these challenges, we have implemented a general AO-driven CASSCF optimizer[178] that provides second order convergence and which may easily be combined with a wide variety of external variational solvers, including DMRG, FCIQMC and their state-averaged solvers. Only the 2-particle density matrix and Hamiltonian integrals are communicated between the CASSCF driver and the external CI solver. Further, the AO-driven algorithm has a low memory and I/O footprint. The current implementation supports calculations with 3000 basis functions and 30–50 active orbitals on a single SMP node with 128 GB memory, without any approximations to the AO integrals.

A simple interface is provided to use an external solver in multiconfigurational calculations. Figure 10 shows how to perform a DMRG-CASSCF calculation by replacing the \texttt{fcisolver} attribute of the CASSCF method. DMRG-SC-NEVPT2[175], and ic-MPS-PT2 and ic-MPS-LCC[181] methods are also available through the interface to the DMRG program package \texttt{Block}[176, 182–184], and the \texttt{ic-MPS-LCC} program of Sharma [181].

### 7.2.4 Molecular properties

At the present stage, the program can compute molecular properties such as analytic nuclear gradients, analytic nuclear Hessians, and NMR shielding parameters at the SCF level. The CCSD and CCSD(T) modules include solvers for the A-equations. As a result, we also provide one-particle and two-particle density matrices, as well as the analytic nuclear gradients, for the CCSD and CCSD(T) methods[167].

For excited states, time-dependent HF (TDHF) and time-dependent DFT (TDDFT) are implemented on top of the SCF module. The relevant analytic nuclear gradients are also programmed[185]. The CCSD module offers another option to obtain excited states using the EOM-IP/EA/EE-CCSD methods. The third option to obtain excited
states is through the multi-root CASCI/CASSCF solvers, optionally followed by the MRPT tool chain. Starting from the multi-root CASCI/CASSCF solutions, the program can compute the density matrices of all the states and the transition density matrices between any two states. One can contract these density matrices with specific AO integrals to obtain different first order molecular properties.

### 7.2.5 Relativistic effects

Many different relativistic treatments are available in PySCF. Scalar relativistic effects can be added to all SCF and post-SCF methods through relativistic effective core potentials (ECP)[186] or the all-electron spin-free X2C[187] relativistic correction. For a more advanced treatment, PySCF also provides 4-component relativistic Hartree-Fock and no-pair MP2 methods with Dirac-Coulomb, Dirac-Coulomb-Gaunt, and Dirac-Coulomb-Breit Hamiltonians. Although not programmed as a standalone module, the no-pair CCSD electron correlation energy can also be computed with the straightforward spin-orbital version of the CCSD program. Using the 4-component Hamiltonian, molecular properties including analytic nuclear gradients and NMR shielding parameters are available at the mean-field level[188].

### 7.2.6 Orbital localizer and result analysis

Two classes of orbital localization methods are available in the package. The first emphasizes the atomic character of the basis functions. The relevant localization functions can generate intrinsic atomic orbitals (IAO)[189], natural atomic orbitals (NAO)[190], and meta-Löwdin orbitals[156] based on orbital projection and orthogonalization. With these AO-based local orbitals, charge distributions can be properly assigned to atoms in population analysis[189]. In the PySCF population analysis code, meta-Löwdin orbitals are the default choice.

The second class, represented by Boys-Foster, Edmiston-Ruedenberg, and Pipek-Mezey localization, require minimizing (or maximizing) the dipole, the Coulomb self-energy, or the atomic charges, to obtain the optimal localized orbitals. The localization routines can take arbitrary orthogonal input orbitals and call the CIAH algorithm to
rapidly converge the solution. For example, using 16 CPU cores, it takes 3 minutes to localize 1620 HF unoccupied orbitals for the \( \text{C}_{60} \) molecule using Boys localization.

A common task when analysing the results of an electronic structure calculation is to visualize the orbitals. Although PySCF does not have a visualization tool itself, it provides a module to convert the given orbital coefficients to the molden\[191\] format which can be read and visualized by other software, e.g. Jmol\[192\]. Figure 11 is an example to run Boys localization for the \( \text{C}_{60} \) HF occupied orbitals and to generate the orbital surfaces of the localized \( \sigma \)-bond orbital in a single Python script.

### 7.2.7 Extended systems with periodic boundary conditions

PBC implementations typically use either plane waves\[4, 193–195\] or local atomic functions\[1, 9, 155, 196–198\] as the underlying orbital basis. The PBC implementation in PySCF uses the local basis formulation, specifically crystalline orbital Gaussian basis functions \( \phi \), expanded in terms of a lattice sum over local Gaussians \( \chi \)

\[
\phi_{k, \chi}(r) = \sum_{T} e^{i\mathbf{k} \cdot \mathbf{T}} \chi(r - \mathbf{T})
\]

where \( \mathbf{k} \) is a vector in the first Brillouin zone and \( \mathbf{T} \) is a lattice translational vector. We use a pure Gaussian basis in our PBC implementation for two reasons: to simplify the development of post-mean-field methods for extended systems and to have a seamless interface and direct comparability to finite-sized quantum chemistry calculations. Local bases are favourable for post-mean-field methods because they are generally quite compact, resulting in small virtual spaces \[199\], and further allow locality to be exploited.

Due to the use of local bases, various boundary conditions can be easily applied in the PBC module, from zero-dimensional systems (molecules) to extended one-, two- and three-dimensional periodic systems.

The PBC module supports both all-electron and pseudopotential calculations. Both separable pseudopotentials (e.g. Goedecker-Teter-Hutter (GTH) pseudopotentials \[46, 47\]) and non-separable pseudopotentials (quantum chemistry ECPs and Burkatzki-Filippi-Dolg pseudopotentials\[200\]) can be used. In the separable pseudopotential
implementation, the associated orbitals and densities are guaranteed to be smooth, allowing a grid-based treatment that uses discrete fast Fourier transforms \([1, 201]\). In both the pseudopotential and all-electron PBC calculations, Coulomb-based integrals are handled via density fitting as described in subsection 7.2.10.

The PBC implementation is organized in direct correspondence to the molecular implementation. We implemented the same function interfaces as in the molecular code, with analogous module and function names. Consequently, methods defined in the molecular part of the code can be seamlessly mixed with the PBC functions without modification, especially in \(\Gamma\)-point calculations where the PBC wave functions are real. Thus, starting from PBC \(\Gamma\)-point mean-field orbitals, one can, for example, carry out CCSD, CASSCF, TDDFT, etc. calculations using the molecular implementations. We also introduce specializations of the PBC methods to support \(k\)-point (Brillouin zone) sampling. The \(k\)-point methods slightly modify the \(\Gamma\)-point data structures, but inherit from and reuse almost all of the \(\Gamma\)-point functionality. Explicit \(k\)-point sampling is supported at the HF and DFT level, and on top of this we have also implemented \(k\)-point MP2, CCSD, CCSD(T) and EOM-CCSD methods[201], with optimizations to carefully distribute work and data across cores. On 100 computational cores, mean-field simulations including unit cells with over 100 atoms, or \(k\)-point CCSD calculations with over 3000 orbitals, can be executed without difficulty.

7.2.8 General AO integral evaluator and J/K builds

Integral evaluation forms the foundation of Gaussian-based electronic structure simulation. The general integral evaluator library LibCINT supports a wide range of GTO integrals, and PySCF exposes simple APIs to access the LibCINT integral functions. As the examples in Figure 12 show, the PySCF integral API allows users to access AO integrals either in a giant array or in individual shells with a single line of Python code. The integrals provided include,

- integrals in the basis of Cartesian, real-spherical and \(j\)-adapted spinor GTOs;
- arbitrary integral expressions built from \(r\), \(p\), and \(\sigma\) polynomials;
• 2-center, 3-center and 4-center 2-electron integrals for the Coulomb operator $1/r_{12}$, range-separated Coulomb operator $\text{erf}(\omega r_{12})/r_{12}$, Gaunt interaction, and Breit interaction.

Using the general AO integral evaluator, the package provides a general AO-driven J/K contraction function. J/K-matrix construction involves a contraction over a high order tensor (e.g. 4-index 2-electron integrals $(ij|kl)$) and a low order tensor (e.g. the 2-index density matrix $\gamma$)

$$J_{ij} = \sum_{kl} (ij|kl) \gamma_{kl}$$
$$K_{il} = \sum_{jk} (ij|kl) \gamma_{jk}$$

When both tensors can be held in memory, the Numpy package offers a convenient tensor contraction function `einsum` to quickly construct J/K matrices. However, it is common for the high order tensor to be too large to fit into the available memory. Using the Einstein summation notation of the Numpy `einsum` function, our AO-driven J/K contraction implementation offers the capability to contract the high order tensor (e.g. 2-electron integrals or their high order derivatives) with multiple density matrices, with a small memory footprint. The J/K contraction function also supports subsystem contraction, in which the 4 indices of the 2-electron integrals are distributed over different segments of the system which may or may not overlap with each other. This subsystem contraction is particularly useful in two scenarios: in fragment-based methods, where the evaluation of Coulomb or exchange energies involves integral contraction over different fragments, and in parallel algorithms, where one partitions the J/K contraction into small segments and distributes them to different computing nodes.

7.2.9 General integral transformations

Integral transformations are another fundamental operation found in quantum chemistry programs. A common kind of integral transformation is to transform the 4 indices of the 2-electron integrals by 4 sets of different orbitals. To satisfy this need, we de-
signed a general integral transformation function to handle the arbitrary AO integrals provided by the LibCint library and arbitrary kinds of orbitals. To reduce disk usage, we use permutation symmetry over $i$ and $j$, $k$ and $l$ in $(ij|kl)$ whenever possible for real integrals.

Integral transformations involve high computational and I/O costs. A standard approach to reduce these costs involves precomputation to reduce integral costs and data compression to increase I/O throughput. However, we have not adopted such an optimization strategy in our implementation because it is against the objective of simplicity for the PySCF package. In our implementation, initialization is not required for the general integral transformation function. Similarly to the AO integral API, the integral transformation can thus be launched with one line of Python code. In the integral data structure, we store the transformed integrals by chunks in the HDF5 format without compression. This choice has two advantages. First, it allows for fast indexing and hyperslab selection for subblocks of the integral array. Second, the integral data can be easily accessed by other program packages without any overhead for parsing the integral storage protocol.

7.2.10 Density fitting

The density fitting (DF) technique is implemented for both finite-sized systems and crystalline systems with periodic boundary conditions.

In finite-sized systems, one can use DF to approximate the J/K matrix and the MO integrals for the HF, DFT and MP2 methods. To improve the performance of the CIAH algorithm, one can use the DF orbital Hessian in the CIAH orbital optimization for Edmiston-Ruedenberg localization and for the HF, DFT and CASSCF algorithms.

In the PBC module, the 2-electron integrals are represented as the product of two 3-index tensors which are treated as DF objects. Based on the requirements of the system being modelled, we have developed various DF representations. When the calculation involves only smooth bases (typically with pseudopotentials), plane waves are used as the auxiliary fitting functions and the DF 3-index tensor is computed within a grid-
based treatment using discrete fast Fourier transforms [201]. When high accuracy in all-electron calculations is required, a mixed density fitting technique is invoked in which the fitting functions are Gaussian functions plus plane waves. Besides the choice of fitting basis, different metrics (e.g. overlap, kinetic, or Coulomb) can be used in the fitting to balance performance against computational accuracy.

The 3-index DF tensor is stored as a giant array in the HDF5 format without compression. With this design, it is straightforward to access the 2-electron integrals with the functions of the PySCF package. Moreover, it allows us to supply 2-electron integrals to calculations by overloading the DF object in cases where direct storage of the 4-index integrals in memory or on disk is infeasible (see discussion in subsection 7.2.11).

7.2.11 Custom Hamiltonians

Most quantum chemistry approximations are not tied to the details of the ab initio molecular or periodic Hamiltonian. This means that they can also be used with arbitrary model Hamiltonians, which is of interest for semi-empirical quantum chemistry calculations as well as condensed-matter model studies. In PySCF, overwriting the predefined Hamiltonian is straightforward. The Hamiltonian is an attribute of the mean-field calculation object. Once the 1-particle and 2-particle integral attributes of the mean-field object are defined, they are used by the mean-field calculation and all subsequent post-Hartree-Fock correlation treatments. Users can thus carry out correlated calculations with model Hamiltonians in exactly the same way as with standard ab initio Hamiltonians. Figure 13 displays an example of how to input a model Hamiltonian.

7.2.12 Interfaces to external programs

PySCF can be used either as the driver to execute external programs or as an independent solver to use as part of a computational workflow involving other software. In PySCF, the DMRG programs BLOCK[176] and CHEMPS2[177, 202] and the FCIQMC program NECI[203, 204] can be used as a replacement for the FCI routine for large
active spaces in the CASCI/CASSCF solver. In the QM/MM interface, by supplying
the charges and the positions of the MM atoms, one can compute the HF, DFT, MP2,
CC, CI and MCSCF energies and their analytic nuclear gradients.

To communicate with other quantum chemistry programs, we provide utility func-
tions to read and write Hamiltonians in the Molpro[205] FCIDUMP format, and ar-
bitrary orbitals in the molden[191] format. The program also supports to write SCF
solution and CI wavefunction in the GAMESS[206] WFN format and to read orbitals
from Molpro XML output. The real space electron density can be output on cubic
grids in the Gaussian[207] cube format.

7.2.13 Numerical tools

Although the Numpy and Scipy libraries provide a wide range of numerical tools
for scientific computing, there are some numerical components commonly found in
quantum chemistry algorithms that are not provided by these libraries. For example,
the direct inversion of the iterative space (DIIS) method[208, 209] is one of the most
commonly used tools in quantum chemistry to speed up optimizations when a second
order algorithm is not available. In PySCF we provide a general DIIS handler for
an object array of arbitrary size and arbitrary data type. In the current implementa-
tion, it supports DIIS optimization both with or without supplying the error vectors.
For the latter case, the differences between the arrays of adjacent iterations are mini-
mized. Large scale eigenvalue and linear equation solvers are also common components
of many quantum chemistry methods. The Davidson diagonalization algorithm and
Arnoldi/Krylov subspace solver are accessible in PySCF through simple APIs.

7.3 Design and implementation of PySCF

While we have tried to provide rich functionality for quantum chemical simulations with
the built-in functions of the PySCF package, it will nonetheless often be the case that
a user’s needs are not covered by the built-in functionality. A major design goal has
been to implement PySCF in a sufficiently flexible way so that users can easily extend
its functionality. To provide robust components for complex problems and non-trivial
workflows, we have made the following general design choices in PySCF:

1. **Language**: Mostly Python, with a little C. We believe that it is easiest to develop and test new functionality in Python. For this reason, most functions in PySCF are written in pure Python. Only a few computational hot spots have been rewritten and optimized in C.

2. **Style**: Mostly functional, with a little object-oriented programming (OOP). Although OOP is a successful and widely used programming paradigm, we feel that it is hard for users to customize typical OOP programs without learning details of the object hierarchy and interfaces. We have adopted a functional programming style, where most functions are pure, and thus can be invoked alone and independently of each other. This allows users to mix functionality with a minimal knowledge of the PySCF internals.

We elaborate on these choices below.

### 7.3.1 Input language

Almost every quantum chemistry package today uses its own custom input language. This is a burden to the user, who must become familiar with a new domain-specific language for every new package. In contrast, PySCF does not have an input language. Rather, the functionality is simply called from an input script written in the host Python language. This choice has clear benefits:

1. *There is no need to learn a domain-specific language.* Python, as a general programming language, is already widely used for numerical computing, and is taught in modern computer science courses. For novices, the language is easy to learn and help is readily available from the large Python community.

2. *One can use all Python language features in the input script.* This allows the input script to implement complex logic and computational workflows, and to carry out tasks (e.g. data processing and plotting) in the same script as the electronic structure simulation (see Figure 14 for an example).
3. *The computational environment is easily extended beyond that provided by the PySCF package.* The PySCF package is a regular Python module which can be mixed and matched with other Python modules to build a personalized computing environment.

4. *Computing can be carried out interactively.* Simulations can be tested, debugged, and executed step by step within the Python interpreter shell.

### 7.3.2 Advancing interactive computing

As discussed above, a strength of the PySCF package is that its functionality can be invoked from the interactive Python shell. However, maximizing its usability in this interactive mode entails additional design optimizations. There are three critical considerations to facilitate such interactive computations:

1. The functions and data need to be easy to access;

2. Functions should be insensitive to execution order (when and how many times a function is called should not affect the result);

3. Computations should not cause (significant) halts in the interactive shell.

To address these requirements, we have enforced the following design rules wherever possible in the package:

1. Functions are pure (i.e. state free). This ensures that they are insensitive to execution order;

2. Method objects (classes) only hold results and control parameters;

3. There is no initialization of functions, or at most a short initialization chain;

4. Methods are placed at both the module level and within classes so that the methods and their documentation can be easily accessed by the interactive shell (see Figure 15).
A practical solution to eliminate halting of the interactive shell is to overlap the REPL (read-eval-print-loop) and task execution. Such task parallelism requires the underlying tasks to be independent of each other. Although certain dependence between methods is inevitable, the above design rules greatly reduce function call dependence. Most functions in PySCF can be safely placed in the background using the standard Python threading and multiprocessing libraries.

7.3.3 Methods as plugins

Ease-of-use is the primary design objective of the PySCF package. However, function simplicity and versatility are difficult to balance in the same software framework. To balance readability and complexity, we have implemented only the basic algorithmic features in the main methods, and placed advanced features in additional “plugins”. For instance, the main mean-field module implements only the basic self-consistent loop. Corrections (such as for relativistic effects) are implemented in an independent plugin module, which can be activated by reassigning the mean-field 1-electron Hamiltonian method at runtime. Although this design increases the complexity of implementation of the plugin functions, the core methods retain a clear structure and are easy to comprehend. Further, this approach decreases the coupling between different features: for example, independent features can be modified and tested independently and combined in calculations. In the package, this plugin design has been widely used, for example, to enable molecular point group symmetry, relativistic corrections, solvation effects, density fitting approximations, the use of second-order orbital optimization, different variational active space solvers, and many other features (Figure 16).

7.3.4 Seamless MPI functionality

The Message Passing Interface (MPI) is the most popular parallel protocol in high performance computing. Although MPI can enable high efficiency parallelism, it is challenging to develop MPI programs that are both simple and efficient. In compiled languages, such programs must explicitly control data communication according to the MPI communication protocol. The most common design is to activate MPI communica-
tion at the beginning of the program and to update the status of the MPI communicator throughout the program. When developing new methods, this often leads to extra effort in code development and debugging.

To retain the simplicity of the PySCF package, we have designed a server-client mechanism to execute parallel code with MPI. We use MPI to start the Python interpreter as a daemon that receives both the functions and data on remote nodes. When a parallel session is activated, the master process sends the functions and the data to the daemons. The function object is decoded remotely and then executed.

Although the detailed implementation of this parallelism strategy is beyond the scope of this paper, a comparison to the traditional MPI program structure suffices to understand the main features of this design.

- In a traditional MPI program, functions and data are treated separately. Parallelism is hard-coded in the functions and communication is performed for the data only. In our parallel implementation, functions and data are communicated on the same footing. This treatment is particularly suitable for dynamic programming languages such as Python, because Python functions are just a type of data which can be serialized at runtime.

- The traditional MPI program runs within a parallel context from the start of the program. The parallel stacks thus have to be maintained throughout the execution period. In our server-client parallel design, the program driver is running in serial mode and the parallel sessions are issued by the driver. The parallel stacks are released when the parallel sessions are finished. Although there is an overhead cost when encoding and decoding the function objects, this overhead is negligible. For example, in our tests the time spent encoding and decoding when starting a parallel session is around 0.1 milliseconds.

- Unlike a traditional MPI program which allows messages to be freely communicated between different threads of different processes, the thread-MPI hybrid programming model is supported at a medium level of granularity in the PySCF parallel structure. MPI communication is handled in the Python functions using
the MPI4PY library. Workloads are distributed among MPI processes by the load-balance functions which are implemented in Python. On each process, the multi-threading code is implemented in the underlying C library extensions.

In the PySCF package, this parallel design allows one to develop code mainly in serial mode and to switch to the MPI mode only when intensive parallelization is required. Figure 17 shows an example to perform a periodic calculation with and without a parallel session. Comparing to the serial mode invocation, we see that the user only has to change the density fitting object to acquire parallel functionality.

7.4 Conclusions

Python and its large collection of third party libraries are helping to revolutionize how we carry out and implement numerical simulations. It is potentially much more productive to solve computational problems within the Python ecosystem because it frees researchers to work at the highest level of abstraction without worrying about the details of complex software implementation. To bring all the benefits of the Python ecosystem to quantum chemistry and electronic structure simulations, we have started the open-source PySCF project.

PySCF is a simple, lightweight, and efficient computational chemistry program package, which supports ab initio calculations for both molecular and extended systems. The package serves as an extensible electronic structure toolbox, providing a large number of fundamental operations with simple APIs to manipulate methods, integrals, and wave functions. We have invested significant effort to ensure simplicity of use and implementation while preserving competitive functionality and performance. We believe that this package contains a new style of program and library design that represents one of the possible paths for future software developments in the field.
Table 3: Features of the PySCF package as of the 1.3 release.

<table>
<thead>
<tr>
<th>Method</th>
<th>Molecule</th>
<th>Solids</th>
<th>Comments</th>
</tr>
</thead>
<tbody>
<tr>
<td>HF</td>
<td>Yes</td>
<td>Yes</td>
<td>(\sim 5000) AOs(^b)</td>
</tr>
<tr>
<td>MP2</td>
<td>Yes</td>
<td>Yes</td>
<td>(\sim 1500) MOs(^b)</td>
</tr>
<tr>
<td>CCSD</td>
<td>Yes</td>
<td>Yes</td>
<td>(\sim 1500) MOs(^b)</td>
</tr>
<tr>
<td>EOM-CCSD</td>
<td>Yes</td>
<td>Yes</td>
<td>(\sim 1500) MOs(^b)</td>
</tr>
<tr>
<td>CCSD(T)</td>
<td>Yes</td>
<td>Yes(^a)</td>
<td>(\sim 1500) MOs(^b)</td>
</tr>
<tr>
<td>MCSCF</td>
<td>Yes</td>
<td>Yes(^a)</td>
<td>(\sim 3000) AOs(^b) 30-50 active orbitals(^c)</td>
</tr>
<tr>
<td>MRPT</td>
<td>Yes</td>
<td>Yes(^a)</td>
<td>(\sim 1500) MOs(^b) 30-50 active orbitals(^c)</td>
</tr>
<tr>
<td>DFT</td>
<td>Yes</td>
<td>Yes</td>
<td>(\sim 5000) AOs(^b)</td>
</tr>
<tr>
<td>TDDFT</td>
<td>Yes</td>
<td>Yes(^a)</td>
<td>(\sim 5000) AOs(^b)</td>
</tr>
<tr>
<td>CISD</td>
<td>Yes</td>
<td>Yes(^a)</td>
<td>(\sim 1500) MOs(^b)</td>
</tr>
<tr>
<td>FCI</td>
<td>Yes</td>
<td>Yes(^a)</td>
<td>(\sim (18e, 18o))(^b)</td>
</tr>
<tr>
<td>Localizer</td>
<td>Yes</td>
<td>No</td>
<td>IAO, NAO, meta-Löwdin Boys, Edmiston-Ruedenberg, Pipek-Meze(^\text{y})</td>
</tr>
<tr>
<td>Relativity</td>
<td>Yes</td>
<td>No</td>
<td>ECP and scalar-relativistic corrections for all methods 2-component, 4-component methods for HF and MP2</td>
</tr>
<tr>
<td>Gradients</td>
<td>Yes</td>
<td>No</td>
<td>HF, DFT, CCSD, CCSD(T), TDDFT</td>
</tr>
<tr>
<td>Hessian</td>
<td>Yes</td>
<td>No</td>
<td>HF and DFT</td>
</tr>
<tr>
<td>Properties</td>
<td>Yes</td>
<td>No</td>
<td>non-relativistic, 4-component relativistic NMR</td>
</tr>
<tr>
<td>Symmetry</td>
<td>Yes</td>
<td>No</td>
<td>D(_{2h}) and subgroup</td>
</tr>
<tr>
<td>AO, MO integrals</td>
<td>Yes</td>
<td>Yes</td>
<td>1-electron, 2-electron integrals</td>
</tr>
<tr>
<td>Density fitting</td>
<td>Yes</td>
<td>Yes</td>
<td>HF, DFT, MP2</td>
</tr>
</tbody>
</table>

\(^{a}\) Only available for \(\Gamma\)-point calculations;
\(^{b}\) An estimation based on a single SMP node with 128 GB memory without density fitting;
\(^{c}\) Using an external DMRG or FCIQMC program as active space solver.
from pyscf import gto, dft
mol = gto.Mole(atom='N 0 0 0; N 0 0 1.1', basis='ccpvtz')
mf = dft.RKS(mol)
mf.xc = '0.2*HF + 0.08*LDA + 0.72*B88, 0.81*LYP + 0.19*VWN'
mf.kernel()

Figure 9: Example to define a custom exchange-correlation functional for a DFT calculation.

from pyscf import gto, scf, mcscf, dmrgscf
mol = gto.Mole(atom='N 0 0 0; N 0 0 1.1', basis='ccpvtz')
mf = scf.RHF(mol).run()
mc = mcscf.CASSCF(mf, 8, 10)  # 8o, 10e
mc.fcisolver = dmrgscf.DMRGCI(mol)
mc.kernel()

Figure 10: Example to enable the DMRG solver in a CASSCF calculation.

from pyscf import gto, scf, lo, tools
mol = gto.Mole(atom=open('c60.xyz').read(),
               basis='ccpvtz')
mf = scf.RHF(mol).run()
orb = lo.Boys(mol).kernel(mf.mo_coeff[:, :180])
tools.molden.from_mo(mol, 'c60.molden', orb)

# Invoke Jmol to plot the orbitals
with open('c60.spt', 'w') as f:
    f.write('load c60.molden; isoSurface MO 002;\n')
import os
os.system('jmol c60.spt')

Figure 11: Example to generate localized orbitals and to plot them in Jmol.
from pyscf import gto

mol = gto.Mole(atom='N 0 0 0; N 0 0 1.1', basis='ccpvtz')
a = mol.intor('cint1e_nuc_sph')  # nuclear attraction as a giant array
a = mol.intor('cint2e_sph')      # 2e integrals as a giant array
a = mol.intor_by_shell('cint2e_sph', (0,0,0))  # (ss|ss) of first N atom

Figure 12: Example to access AO integrals.

# 10-site Hubbard model at half-filling with U/t = 4
import numpy as np
from pyscf import gto, scf, ao2mo, cc

cmpmol = gto.Mole(verbose=4)
ncmpmol.nelectron = n = 10
t, u = 1., 4.

mf = scf.RHF(mol)
h1 = np.zeros((n,n))
for i in range(n-1):
    h1[i,i+1] = h1[i+1,i] = t
mf.get_hcore = lambda *args: h1
mf.get_ovlp = lambda *args: np.eye(n)
mf._eri = np.zeros((n,n,n,n))
for i in range(n):
    mf._eri[i,i,i,i] = u

# 2e Hamiltonian in 4-fold symmetry
mf._eri = ao2mo.restore(4, mf._eri, n)
mf.run()
cc.CCSD(mf).run()

Figure 13: Example to use a custom Hamiltonian.
```python
import numpy as np
from pyscf import gto, scf
bond = np.arange(0.8, 5.0, .1)
dm_init = None
e_hf = []
for r in reversed(bond):
    mol = gto.Mole(atom=[['N', 0, 0, 0], ['N', 0, 0, r]],
                   basis='ccpvtz')
    mf = scf.RHF(mol).run(dm_init)
    dm_init = mf.make_rdm1()
    e_hf.append(mf.e_tot)

from matplotlib import pyplot
pyplot.plot(bond, e_hf[::-1])
pyplot.show()
```

Figure 14: Using Python to combine the calculation and data post-processing in one script.

```python
mol = gto.Mole(atom='N 0 0 0; N 0 0 1.1', basis='ccpvtz')
mf = scf.newton(scf.sfx2c(scf.density_fit(scf.RHF(mol)))).run()
```

Figure 15: Accessing documentation within the IPython shell. The question mark activates the documentation window in the bottom area. The pop-up menu for code auto-completion is triggered by the <Tab> key.

```python
from pyscf import gto, scf
```

Figure 16: Example to use plugins in PySCF. The mean-field calculation is decorated by the density fitting approximation, X2C relativistic correction and second order SCF solver.
Serial mode
# run in cmdline:
# python input.py
from pyscf.pbc import gto, scf
from pyscf.pbc import df
cell = gto.Cell()
cell.atom = 'H 0 0; H 0 0 0.7'
cell.basis = 'ccpvdz'
# unit cell lattice vectors
cell.a = '2 0 0; 0 2 0; 0 0 2'
# grid for numerical integration
cell.gs = [10,10,10]
mf = scf.RHF(cell)
mf.with_df = df.DF(cell)
mf.kernel()

MPI mode
# run in cmdline:
# mpirun -np 4 python input.py
from pyscf.pbc import gto, scf
from mpi4pyscf.pbc import df
cell = gto.Cell()
cell.atom = 'H 0 0; H 0 0 0.7'
cell.basis = 'ccpvdz'
# unit cell lattice vectors
cell.a = '2 0 0; 0 2 0; 0 0 2'
# grid for numerical integration
cell.gs = [10,10,10]
mf = scf.RHF(cell)
mf.with_df = df.DF(cell)
mf.kernel()

Figure 17: Comparison of the input script for serial-mode and MPI-mode calculations. Except for the module to import, the MPI parallel mode takes exactly the same input as the serial mode.
8 Future Directions

The current and future goal is achieving chemical accuracy in solids. In the previous work on solids there were various sources of errors, inherent in any periodic given calculation, that need to be carefully handled for future benchmark calculations. Some of the errors include the error due to finite-size effects and finally use of finite basis set and inexact theory. In this final section we will discuss some of the preliminary work we’ve done in trying to achieve the goals of reducing finite-size effects and our use of EOM-CCSD for excited states. Here we will look into the EOM-CCSD* methods and possible improvements of those methods to achieve chemical accuracy.

8.1 Finite-size Errors

The finite-size effects is the error from approximating a thermodynamic limit, found from an infinite number of k-points, with only a finite number of k-points. The quality of the system properties obtained in a calculation is related to the uncertainty of the extrapolation to this thermodynamic limit. Figure 6 and 7 show that there is some room for improvement in respect to improving finite-size errors. In our previous discussion of silicon and diamond systems we mentioned that the leading source of error was the finite-size errors of the Hartree-Fock eigenvalues. Using a converged density, $\rho_\infty(\mathbf{r})$,
one can then use this to form the Fock matrix at a given k-point \( F_k \), and use these as the one-electron matrix elements for a given correlated theory.

Since our publication on the excited states of diamond and silicon [201], finite-size effects in correlated methods such as MP2 and CCSD have been studied [124]. It still remains for us to investigate the MP2 and CCSD scalings using the converged Hartree-Fock eigenvalues at the infinite k-point limit as reported in our earlier work [201]. Further testing on other systems needs to be done in order to estimate the true benefit this method offers both for MP2 and perhaps other correlated methods.

8.2 Periodic EOM-CCSD*

We have implemented the EOM-CCSD* as described earlier and did the necessary work to implement the momentum symmetry for application to periodic systems. The current implementation is fairly scalable and has been run on a small subset of systems and basis sets of those considered in our previous work on the semiconductors diamond and silicon. Results for some modest system sizes are shown in Figure 18. We see that the results for EOM-CCSD* systematically underestimate the band gap compared to EOM-CCSD, worsening performance. We also note that the band gap lowers with respect to increasing basis set; at the complete basis set limit we would expect the band gap to be slightly less than those shown here, assuming the extrapolations continue their trend with respect to increasing number of k-points. While underestimating the band gap at DZVP, the results should worsen, if the same trend holds, when moving towards the basis set limit. While the results were discouraging, we then investigated the EOM-CCSD* on a variety of molecular systems in order to test the performance of the IP and EA equation-of-motion variants. We note that while the IP had been tested before, this represents the first time the EA EOM-CCSD* has been properly benchmarked on a series of molecular systems.

Because of the poor performance and lack of benchmark data I ran the EOM-CCSD* methods on a subset of the GW100 set, a dataset composed of CCSD(T) reference data for ionization potentials and electron affinities for molecular systems [210]. We see this
Figure 18: The band gap for diamond calculated at the EOM-CCSD level of theory for DZVP and TZVP with varying number of k-points is shown in blue and red, respectively. Also shown is the DZVP calculated at the EOM-CCSD* level of theory, denoted by the pink line.

in Figure 19, where we show the performance of the IP, EA, and calculated band gap against the reference data. Here we see that the EOM-CCSD fairs slightly better on average compared to the EOM-CCSD*. The average of the EOM-CCSD calculations agree quite well the the reference data for CCSD(T) and agrees better than the average of the EOM-CCSD*. The spread of the EOM-CCSD and EOM-CCSD* are quite similar for the IP, but we see a worse spread for the EOM-CCSD* in the EA and resulting calculation of the band gap. From this data, there doesn’t seem to be any particular reason to use the perturbative method over the vanilla EOM-CCSD. It was checked that the outliers for the two methods were on the same molecules, so it isn’t necessary that the perturbative equation-of-motion behaves better on a certain subset of the data.

The general trend we see is that the EOM-CCSD* underestimates the ionization potential by roughly 0.15 eV and overestimates the electron affinity by about 0.10 eV, resulting in a band gap that is on average 0.25 eV too small. This agrees quite well with what we saw in Figure 18, where our EOM-CCSD* underestimated band gaps compared
with those at the EOM-CCSD level of theory. As we’ve seen, the conclusion for both the molecular and periodic systems both behave similarly in their underestimation of the band gap. It is reasonable then to conclude that when benchmarking perturbative EOM-CCSD methods for molecular systems, perturbative methods that perform well in the molecular case should behave similarly in the periodic case.

Since EOM-CCSD works well, it would be beneficial to improve upon it, yet steer away from the high cost of inclusion of full triples in EOM-CCSDT. Because of this, we have decided to build upon the EOM-CCSD* framework to improve excitation energies, particularly in the case of the electron affinity, to move us closer to chemical accuracy. Starting with the full EOM-CCSD* equations and inclusion of triples to perturb the $T_1$ and $T_2$ amplitudes as done by Matthews and Stanton [39], our hope is to learn from the reference data what diagrams are necessary to include in order to obtain accurate energies like those from CCSD(T). It is then our hope that the models that perform well for the GW100 set will then translate into better performing excited state methods for solids. Our next section will develop these methodologies further.

### 8.3 Beyond EOM-CCSD*

The unsaid goal of creating a perturbative triples for excited states is to provide excited state energies with the same chemical accuracy as in ground state CCSD(T). The original approach that ultimately produced the CCSD(T) equations had elements of an empirical approach, where a subset of diagrammatic terms were originally included based on perturbation theory and then their final inclusion into the model was based on the model’s performance. Later, Stanton [211] applied a Löwdin partitioning technique, like in the case of EOM-CCSD*, to ultimately provide a fairly reasonable justification for why the CCSD(T) method worked. This knowledge was then used to develop other higher scaling theories with even better accuracies, like the CCSDT(Q) method [212].

**What terms are needed in the perturbation expansion to achieve high accuracy?**

The high accuracy in CCSD(T) was found through empirical observation; including only some of the perturbation expansion at a given order was found to give better accuracy
than turning on all of the terms. It is worthwhile to note that there was no particular reasoning for why some terms worked better than others. For instance, the accuracy of the model had little to do with the magnitude of the higher-order diagrams included [211]. In addition to which terms need to be included, there is the question of whether left eigenvectors should be included or approximated by the right eigenvectors. The left lambda amplitudes of course make no contribution to the CCSD(T) theory despite being formally present in the perturbative expansion for CCSD(T). An argument for their exclusion is based on the building of approximate lambda amplitudes from the right amplitudes. Beyond this, there is the possibility for additional approximations. It was found that the inclusion of MP2 amplitudes instead of their CCSD converged \( t_1 \) and \( t_2 \) amplitudes for the problem of core ionization processes [37]. In addition, it would be reasonable to expect that actual triples corrections enter into the equation through its perturbative correction to the \( t_1 \) and \( t_2 \) amplitudes. This latter point has been explored by Matthews and Stanton [39], as well as the idea that a perturbative treatment of the ground state energy might be necessary. This could remove the discrepancy of treating the ground and excited states at different levels of theory and has lead to some improvement in the overall energy.

Our search can be limited down to a smaller number of diagrams if we only consider the turning on and off of so-called skeleton diagrams, where arrows for particle/hole arrows are excluded from the diagram. The thought behind their exclusion is that by treating only a subset of terms coming from the skeleton diagrams you generally have will get a poor answer. In some sense this can be thought of treating the holes and particles on different footings by choosing to keep some of these directional diagrams over others. Here we note also that the CCSD(T) diagrams too are based off of inclusion of whole skeleton diagrams and so we will continue in this way for our study. In equation 40 the last two terms correspond to the same skeleton diagram. In equation 41 the first two and second two terms correspond to the same skeleton diagrams. After that, each term remaining in equation 41 can be paired with a corresponding term coming from the same skeleton diagram in equation 42. In some sense, they have found
that removing complete skeleton diagrams produced better results than removing any individual directional diagrams. This lends credence to the fact that we should only turn some diagrams on and off together, which reduces the size of the overall search space.

8.4 Conclusion

While our EOM-CCSD has performed well in both model metallic systems and the semiconductors of diamond and silicon, there is still room to improve in terms of reducing finite-size effects and achieving higher accuracy through perturbative methods. We explored ways of reducing these finite-size effects through using converged Hartree-Fock eigenvalues, a result that already shows promise in our limited test cases. Likewise, we have looked at some existing perturbative expansions for excited states. These results suggest that there is a need to improve on these perturbative expansions through our benchmarks against the $GW_{100}$ molecular dataset.

The EOM-CCSD method has already been shown to do as well as the existing methods for treating excited states in periodic systems. The next step is to benchmark the EOM-CCSD method for a variety of semiconductors and compare against the existing DFT and $GW$ methods. For this, the reduction of finite-size effects will be essential to improve convergence to the thermodynamic limit, reducing both uncertainty in our extrapolation measurement and the size of the k-point mesh needed to reach this limit. Although not discussed in detail here, other improvements need to be made even at the mean-field level to achieve the complete basis set limit for periodic calculations as right now issues with linear dependency arise in solving the Hartree-Fock equations. Finally, we can explore the ways in which to improve the current perturbative methods for excited states as described in detail above. With these improvements stated here, the periodic EOM-CCSD method is a promising candidate for becoming the standard in calculating excited states in solids.
References


127


(143) Nooijen, M.; Snijders, J. G. Diagrammatic analysis and application of the
coupled cluster response approach to ground-state expectation values. *Int. J. 


(145) Jones, E.; Oliphant, T.; Peterson, P. {SciPy}: Open source scientific tools for 
{Python}. *2014*.

(146) Dalcin, L. D.; Paz, R. R.; Kler, P. A.; Cosimo, A. Parallel distributed computing 

(147) Turney, J. M.; Simmonett, A. C.; Parrish, R. M.; Hohenstein, E. G.; Evangelista, F. A.; 

(148) Bahn, S. R.; Jacobsen, K. W. An object-oriented scripting interface to a legacy 

Gunter, D.; Chevrier, V. L.; Persson, K. A.; Ceder, G. Python Materials Ge-

(150) Jacob, C. R.; Beyhan, S. M.; Bulo, R. E.; Gomes, A. S. P.; Götz, A. W.; 
Kiewisch, K.; Sikkema, J.; Visscher, L. PyADF—A scripting framework for 

(151) Hirata, S. Tensor contraction engine: Abstraction and automated parallel im-
plementation of configuration-interaction, coupled-cluster, and many-body per-

(152) Neuscamman, E.; Yanai, T.; Chan, G. K.-L. Quadratic canonical transformation 


(178) Sun, Q.; Yang, J.; Chan, G. K.-L. A general second order complete active space self-consistent-field solver for large-scale systems. *Chem. Phys. Lett.* **2017**.


Figure 19: Performance of EOM-CCSD* against EOM-CCSD for a subset of the GW100 molecular dataset for calculation of (top) ionization potential, (middle) electron affinity, and (bottom) band gaps versus CCSD(T) reference data.