QUANTUM MONTE CARLO DEVELOPMENTS FOR DISCRETE AND CONTINUOUS SPACES

A Dissertation

Presented to the Faculty of the Graduate School

of Cornell University

in Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

by

Frank Robert Petruzielo Jr

May 2012

© 2012 Frank Robert Petruzielo Jr

ALL RIGHTS RESERVED

QUANTUM MONTE CARLO DEVELOPMENTS FOR DISCRETE AND CONTINUOUS SPACES

Frank Robert Petruzielo Jr, Ph.D.

Cornell University 2012

This thesis details four research projects related to zero temperature quantum Monte Carlo. Chapters 2-4 focus on continuum quantum Monte Carlo, and specifically its application to molecular systems; whereas Chapter 5 focuses on quantum Monte Carlo in a discrete space.

Chapter 2 focuses on improving upon the single-particle basis functions employed in quantum Monte Carlo calculations for molecular systems. For calculations requiring non-diverging pseudopotentials, a class of functions is introduced that is capable of producing the short- and long-range asymptotic behavior of the exact wavefunction. It is demonstrated that this form of basis function produces superior accuracy and efficiency when compared to the basis sets typically employed in quantum Monte Carlo.

Although the basis functions introduced in Chapter 2 are capable of producing superior results, it is necessary that the parameters of the functional form are near-optimal for the full potential of the functions to be realized. Chapter 3 introduces a simple yet general method for constructing basis sets of a desired functional form appropriate for molecular electronic structure calculations. A standard basis set is created for each of the elements from hydrogen to argon.

Chapter 4 explores the effect of different aspects of the trial wavefunction on the accuracy of quantum Monte Carlo. By systematically testing the effect of the basis size, orbital quality, and determinant expansion quality, this work offers guidance to quantum Monte Carlo practitioners for achieving results to within chemical accuracy of experiment.

In Chapter 5, semistochastic projection, a hybrid of deterministic and stochastic projection, is introduced for finding the dominant eigenvalue and eigenvector of a matrix. This method, like stochastic projection, is applicable to matrices well beyond the size that can be handled by deterministic methods. Semistochastic projection improves over stochastic projection by significantly reducing the computational time required to obtain the eigenvalue within a specified statistical uncertainty. After the semistochastic projection method is introduced, it is applied to determine the ground state energy of the Hamiltonian in a discrete basis. This special case of semistochastic projection, dubbed semistochastic quantum Monte Carlo, is shown to be orders of magnitude more efficient than stochastic quantum Monte Carlo.

BIOGRAPHICAL SKETCH

Frank Petruzielo was born and raised in southern Florida, and he was extremely fortunate to have two loving parents who supported his varied interests in soccer, karate, and science. As a teenager, Frank moved to Atlanta where he graduated first in his class at Cherokee High School. Frank earned a B.S. in physics and mathematics from Berry College in 2006. At Cornell, Frank quickly joined the research group of Prof. Cyrus Umrigar, adjusted to snowy winters, and met his future wife Robin.

I dedicate this dissertation to my loving and supportive family.

ACKNOWLEDGMENTS

I owe an enormous debt to my advisor and mentor Prof. Cyrus Umrigar, who has provided me with an opportunity to do fascinating science. Not only has he shaped the way I conduct research and think about scientific problems, he has pushed me to develop and defend my own ideas. As Dr. Umrigar is always thinking about science, he has been an invaluable resource through many late night and weekend phone calls. His door is always open.

I greatly appreciate my committee members Prof. Tomas Arias, Prof. Garnet Chan, and Prof. Paul McEuen for critically reading this thesis. I thank current Umrigar group members Hitesh Changlani and Adam Holmes for the valuable collaborations. I am especially indebted to Julien Toulouse who has taken great interest in much of my work. I thank Arias group members for shared computing resources, and Duane Loh for many stimulating conversations. I thank Claudia Filippi for valuable discussions.

I thank the following sources for supporting this work financially and through computational resources: NDSEG fellowship, NSF (DMR-0908653, CHE-1004603), DOE (DOE-DE-FG05-08OR23339, CSMN program), Cornell NanoScale Facility and Computation Center for Nanotechnology Innovation at Rensselaer Polytechnic Institute.

Finally, I give my love and thanks to my family. My parents have gone above and beyond to ensure that my life in Ithaca is comfortable. In addition, my parents have been an endless source of encouragement. My sister has been a great voice of moderation: she always lightens the mood! My wife Robin has supported and believed in me every step of the way. As a fellow physics graduate student, she understands the many challenges of pursuing a Ph.D. in physics. She has helped me to become a better scientific writer.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

INTRODUCTION

Quantum Monte Carlo methods [1], described in this thesis, are a class of computational methods which allow one to study physical systems and phenomena that would otherwise be inaccessible due to their size or accuracy required. Having the ability to study these problems comes at a steep price. The results are obtained with a statistical error bar which decays inversely with the square root of the computer time.

If most cases, if an alternative to quantum Monte Carlo is possible for a particular calculation, then that alternative should be used. For instance, density functional theory [2] takes dramatically less time than continuum quantum Monte Carlo methods to compute properties of materials. Similarly, when possible, exact diagonalization wins hands down over quantum Monte Carlo calculations for discrete systems.

When to apply continuum quantum Monte Carlo is not entirely black and white. Continuum quantum Monte Carlo has favorable scaling with system size when compared to other highly accurate wavefunction based methods (Coupled-Cluster Theory, Configuration Interaction, etc. [3]). Additionally, continuum quantum Monte Carlo allows one to work in the infinite basis set limit, whereas basis set extrapolation techniques are required for the other wavefunction based methods. Consequently, as system size increases there is a cross over when quantum Monte Carlo methods become the most efficient option.

The choice between quantum Monte Carlo and deterministic methods is more clear for discrete systems or continuum systems in a discrete basis. In this

case, if better than mean-field accuracy is required, quantum Monte Carlo is applied when Hilbert spaces become too large to be handled deterministically. This is relevant for studying the physics of model systems, like the Hubbard Model, where modest sized lattices quickly reach this threshold.

The fact of the matter is quantum Monte Carlo methods have a wide range of potential applications. Their systematic accuracy has been demonstrated across a wide range of systems: molecules, solids, quantum dots, clusters, and model systems [4]. Additionally, the algorithms are almost embarrassingly parallel and have low memory requirements making them attractive for state of the art supercomputers. For all of these reasons, quantum Monte Carlo methods will continue to play a pivotal role in the fields of computational physics and quantum chemistry.

1.1 Quantum Monte Carlo

Perhaps the most fundamental problem in quantum mechanics is: Given a Hamiltonian $\hat{H}$, find the ground state energy E_0 of the time-independent Schrödinger equation. The ground state wavefunction, $|\psi_0\rangle$, corresponding to E_0, can be expressed in a complete basis as

$$|\psi_0\rangle = \sum_i b_i |\phi_i\rangle, \tag{1.1}$$

where

$$b_i = \langle \phi_i | \psi_0 \rangle. \tag{1.2}$$

For the above expressions and what follows, replace summation with integration for a continuous basis.

A fundamental object of interest in zero temperature quantum Monte Carlo is

$$E[\hat{P}, |\psi_T\rangle] = \frac{\langle\psi_T|\hat{P}\hat{H}|\psi_T\rangle}{\langle\psi_T|\hat{P}|\psi_T\rangle}, \tag{1.3}$$

where the trial wavefunction

$$|\psi_T\rangle = \sum_i c_i|\phi_i\rangle, \tag{1.4}$$

is an approximation to the ground state wavefunction, and $\hat{P}$ is an operator, the choice of which leads to different quantum Monte Carlo theories.

A general quantum Monte Carlo theory for evaluating Eqn. (1.3) can be derived as follows. Expanding Eqn. (1.3) by inserting a complete basis in several places yields

$$\begin{aligned}\frac{\langle\psi_T|\hat{P}\hat{H}|\psi_T\rangle}{\langle\psi_T|\hat{P}|\psi_T\rangle} &= \frac{\sum_{ijk}\langle\psi_T|\phi_i\rangle\langle\phi_i|\hat{P}|\phi_j\rangle\langle\phi_j|\hat{H}|\phi_k\rangle\langle\phi_k|\psi_T\rangle}{\sum_{ij}\langle\psi_T|\phi_i\rangle\langle\phi_i|\hat{P}|\phi_j\rangle\langle\phi_j|\psi_T\rangle} \\ &= \frac{\sum_{ijk} c_i P_{ij} H_{jk} c_k}{\sum_{ij} c_i P_{ij} c_j}.\end{aligned} \tag{1.5}$$

Our goal is to write Eqn. (1.5) as a statistical expectation value with respect to some distribution. The efficiency of evaluating the statistical expectation value depends on the distribution and hence, it is pertinent to introduce the idea of importance sampling which allows for the alteration of the distribution. Importance sampling is accomplished by introducing a guiding wavefunction

$$|\psi_G\rangle = \sum_i d_i|\phi_i\rangle. \tag{1.6}$$

Massaging Eqn. (1.5) yields

$$\begin{aligned}\frac{\sum_{ijk} c_i P_{ij} H_{jk} c_k}{\sum_{ij} c_i P_{ij} c_j} &= \frac{\sum_{ijk} c_i P_{ij} \frac{d_j^2}{d_j^2} H_{jk} c_k}{\sum_{ij} c_i P_{ij} \frac{d_j^2}{d_j^2} c_j} \\ &= \frac{\sum_j d_j^2 \left(\sum_i \frac{c_i P_{ij}}{d_j}\right) \left(\sum_k \frac{H_{jk} c_k}{d_j}\right)}{\sum_j d_j^2 \left(\sum_i \frac{c_i P_{ij}}{d_j}\right) \left(\frac{c_j}{d_j}\right)} \\ &= \frac{\sum_j d_j^2 \left(\sum_i \frac{c_i P_{ij}}{d_j}\right) E_{L,j}}{\sum_j d_j^2 \left(\sum_i \frac{c_i P_{ij}}{d_j}\right) \left(\frac{c_j}{d_j}\right)}, \end{aligned} \tag{1.7}$$

where

$$E_{L,j} = \sum_k \frac{H_{jk} c_k}{d_j}, \tag{1.8}$$

is the local energy. It is assumed that $d_j \neq 0 \; \forall j$.

1.2 Variational Monte Carlo

Variational Monte Carlo is defined by the choice, $\hat{P} = \hat{I}$, the identity operator. By the variational principle,

$$E[\hat{I}, |\psi_T\rangle] = \frac{\langle \psi_T | \hat{H} | \psi_T \rangle}{\langle \psi_T | \psi_T \rangle} \geq E_0, \tag{1.9}$$

with equality holding when $|\psi_T\rangle = |\psi_0\rangle$. To evaluate $E[\hat{I}, |\psi_T\rangle]$, Eqn. (1.7) is used,

$$\begin{aligned} E[\hat{I}, |\psi_T\rangle] &= \frac{\sum_j d_j^2 \left(\sum_i \frac{c_i \delta_{ij}}{d_j}\right) E_{L,j}}{\sum_j d_j^2 \left(\sum_i \frac{c_i \delta_{ij}}{d_j}\right) \left(\frac{c_j}{d_j}\right)} \\ &= \frac{\sum_j \frac{d_j^2}{\sum_l d_l^2} \frac{c_j}{d_j} E_{L,j}}{\sum_j \frac{d_j^2}{\sum_l d_l^2} \frac{c_j^2}{d_j^2}} \\ &= \frac{\sum_j \rho(j) \frac{c_j}{d_j} E_{L,j}}{\sum_j \rho(j) \frac{c_j^2}{d_j^2}}, \end{aligned} \tag{1.10}$$

where

$$\rho(j) = \frac{d_j^2}{\sum_l d_l^2}. \tag{1.11}$$

If the sums/integrals cannot be performed analytically and the number of states (possibly infinite) is to large to store in memory, then Eqn. (1.10) is evaluated using Monte Carlo integration,

$$E[\hat{I}, |\psi_T\rangle] \approx \frac{\sum_{j=1}^{N_{\mathrm{MC}}} \frac{c_j}{d_j} E_{L,j}}{\sum_{j=1}^{N_{\mathrm{MC}}} \frac{c_j^2}{d_j^2}}. \tag{1.12}$$

A description of the Monte Carlo evaluation of 1.12 requires the concept of a stochastic representation of a distribution. Given some probability distribution ρ, a stochastic representation of ρ is a collection of random samples from ρ. An accurate representation of ρ requires a large collection of random samples, typically too large to store in memory. Hence, a natural way to generate this collection of random samples is through a time series. This time series is generated using the Metropolis-Hastings algorithm [5, 6]. A detailed discussion of the Metropolis-Hastings algorithm is given in Appendix A.

Due to the ratio of expectation values in Eqn. (1.10), an unbiased Monte Carlo estimate of $E[\hat{I}, |\psi_T\rangle]$ is nontrivial. In fact, Eqn. (1.12) is a biased estimate of $E[\hat{I}, |\psi_T\rangle]$ when $|\psi_T\rangle \neq |\psi_G\rangle$ for finite N_{MC}. The bias goes to zero as $1/N_{\mathrm{MC}}$. The unbiased expression is more complicated. A detailed discussion of Monte Carlo estimators is given in Appendix A.

When $|\psi_T\rangle = |\psi_0\rangle$, $E_{L,j} = E_0 \frac{c_j}{d_j}$. In this case, for each Monte Carlo sample the numerator and denominator of Eqn. (1.12) each fluctuate, but their ratio has no fluctuations and is equal to E_0.

Finally, variational Monte Carlo has no sign problem since each term in the

numerator and denominator of Eqn. (1.12) depends only on the state sampled.

1.2.1 Importance Sampling with the Trial Wavefunction

Consider the case where the guiding wavefunction is the trial wavefunction,

$$d_i = c_i \neq 0 \; \forall \; i. \tag{1.13}$$

With this choice, Eqn. (1.12) yields

$$E[\hat{I}, |\psi_T\rangle] = \frac{1}{N_{\mathrm{MC}}} \sum_{j=1}^{N_{\mathrm{MC}}} E_{L,j}. \tag{1.14}$$

Importance sampling with the trial wavefunction is the standard choice for variational Monte Carlo. However, for a fixed $|\psi_T\rangle$, $|\psi_G\rangle = |\psi_T\rangle$ may not be the optimal choice. Although there are no fluctuations in the denominator of Eqn. (1.14) (equal to 1 for each Monte Carlo sample), there could be large fluctuations in the numerator.

On the other hand, if $|\psi_G\rangle$ and $|\psi_T\rangle$ are sufficiently different such that $\{c_j/d_j\}$ span a wide range of values, then there will be large fluctuations in Eqn. (1.12).

1.3 Projector Monte Carlo

Projector Monte Carlo is defined by the choice, $\hat{P}|\psi_T\rangle \propto |\psi_0\rangle$. In this case,

$$\begin{aligned} E[\hat{P}, |\psi_T\rangle] &= \frac{\langle \psi_T | \hat{P}\hat{H} | \psi_T \rangle}{\langle \psi_T | \hat{P} | \psi_T \rangle} \\ &= \frac{\langle \psi_0 | \hat{H} | \psi_T \rangle}{\langle \psi_0 | \psi_T \rangle} \\ &= E_0. \end{aligned} \tag{1.15}$$

To evaluate $E[\hat{P}, |\psi_T\rangle]$, Eqn. (1.7) is used,

$$E[\hat{P}, |\psi_T\rangle] = \frac{\sum_j d_j^2 \left(\sum_i \frac{c_i P_{ij}}{d_j}\right) E_{L,j}}{\sum_j d_j^2 \left(\sum_i \frac{c_i P_{ij}}{d_j}\right) \left(\frac{c_j}{d_j}\right)} = \frac{\sum_j b_j d_j E_{L,j}}{\sum_j b_j d_j \frac{c_j}{d_j}}. \tag{1.16}$$

If the projector cannot be applied analytically, then Monte Carlo is used to calculate $\{b_j d_j\}$. Additionally, Monte Carlo is typically used to evaluate the sum over states in Eqn. (1.16), but this is only necessary if the number of states is too large to store in memory.

Like the case of variational Monte Carlo, Monte Carlo evaluation of Eqn. (1.16) requires a time series of random samples. Unlike the case of variational Monte Carlo, $\rho(j) = b_j d_j$ is unknown, so it cannot be sampled with the Metropolis-Hastings algorithm. Instead each random sample in the time series is generated by applying the projector to the previous sample.

In most cases, the projector does not preserve normalization. This makes it necessary for each random sample to both represent a particular state and carry a weight. This state and weight pair is known as a walker. The stochastic representation of ρ is then a time series of walkers. In practice this prescription is inefficient because the log of the walker weights undergo a random walk and some generations contribute much more than others. Monte Carlo averages are most efficient, in the statistical error sense, when walker weights are all the same. Therefore, it is better to have a population of walkers at each step of the time series instead of a single walker. When a walker's weight becomes larger than an upper threshold, it is split into multiple walkers; and, when a walker's weight becomes smaller than a lower threshold, it is combined with other walkers in a manner such that the expectation value is unchanged.

Using the concepts of the previous discussion, the Monte Carlo evaluation of $b_j d_j$ is

$$b_j d_j = \frac{1}{N_{\mathrm{MC}}} \sum_{n=1}^{N_{\mathrm{MC}}} w_j(n), \tag{1.17}$$

where $w_j(n)$ is the total weight of the walkers on state j for the n^{th} step of the time series. Hence, for projector Monte Carlo,

$$\begin{aligned} E[\hat{P}, |\psi_T\rangle] &= \frac{\sum_j d_j b_j E_{L,j}}{\sum_j d_j b_j \frac{c_j}{d_j}} \\ &= \frac{\sum_j \sum_{n=1}^{N_{\mathrm{MC}}} w_j(n) E_{L,j}}{\sum_j \sum_{n=1}^{N_{\mathrm{MC}}} w_j(n) \frac{c_j}{d_j}} \\ &= \frac{\sum_{n=1}^{N_{\mathrm{MC}}} \sum_j w_j(n) E_{L,j}}{\sum_{n=1}^{N_{\mathrm{MC}}} \sum_j w_j(n) \frac{c_j}{d_j}}. \end{aligned} \tag{1.18}$$

Note that even though the projector Monte Carlo energy does not depend on $|\psi_T\rangle$, the fluctuations do depend on $|\psi_T\rangle$.

Unfortunately, projector Monte Carlo has a sign problem because, in general, the sign of $w_j(n)$ depends on n, not just the state j. An efficient implementation of projector Monte Carlo should make the sign of $w_j(n)$ independent of n. This is accomplished by either modifying the projector or how the projector is applied. In general, these modifications introduce a $|\psi_T\rangle$ dependent bias in the Monte Carlo estimate of the energy.

1.4 Wavefunction Optimization

As seen in Eqn. (1.12), the variational Monte Carlo energy and its fluctuations depends intimately on the trial wavefunction. Additionally, the fluctuations of the projector Monte Carlo energy, and in practical implementations, the

energy itself, depend on the trial wavefunction. Therefore, a high quality trial wavefunction is an essential ingredient to quantum Monte Carlo.

What qualifies as a high quality trial wavefunction? Undoubtedly, the functional form should be capable of capturing the relevant physics or chemistry of the problem of interest. Beyond this requirement, specifying the trial wavefunction boils down to an optimization problem.

At the heart of an optimization problem is the objective function, the quantity being optimized. The question of what to optimize may seem obvious since the main quantity of interest is the energy, but the story is far more complex.

1.5 Overview of Thesis

The following chapters detail four research projects related to zero temperature quantum Monte Carlo. Chapters 2-4 focus on continuum quantum Monte Carlo, and specifically its application to molecular systems; whereas Chapter 5 focuses on quantum Monte Carlo in a discrete space.

1.5.1 Overview of Chapter 2

As mentioned previously, a high quality trial wavefunction is an essential ingredient of quantum Monte Carlo. For continuum quantum Monte Carlo applied to molecular systems, the most common form of trial wavefunction employed is the Slater-Jastrow wavefunction, which is a product of a sum of Slater determinants and a Jastrow factor.

A Slater determinant is a determinant of single-particle orbitals, each expanded in a set of single-particle basis functions. By construction, the Slater determinant is antisymmetric under exchange of electrons, which is a minimal requirement for a fermionic wavefunction. The Jastrow factor is a positive function of inter-electronic coordinates. The Jastrow factor introduces explicit particle-particle correlations and is essential for accurate and efficient quantum Monte Carlo calculations.

Chapter 2 focuses on improving a specific aspect of the Slater-Jastrow wavefunction, namely the single-particle basis functions. Molecular calculations in quantum Monte Carlo frequently employ a Gaussian basis, which cannot produce the correct asymptotic behavior of the exact wavefunction. For calculations requiring non-diverging pseudopotentials, Gauss-Slater basis functions, which have the correct short- and long-range asymptotic behavior, are introduced. Gauss-Slater functions behave like Gaussians at short distances and Slaters at long distances. It is demonstrated that this form of basis function produces superior accuracy and efficiency when compared to the basis sets typically employed in quantum Monte Carlo.

1.5.2 Overview of Chapter 3

Chapter 3 expands upon the idea of using Gauss-Slater basis functions in quantum Monte Carlo. Although Gauss-Slater functions are capable of producing superior results, it is necessary that the parameters of the functional form are near-optimal for the full potential of these functions to be realized. These parameters can be optimized for each system, which is done in Chapter 2; but

it is desirable, or in the case of large systems necessary, to have a set of basis functions with near-optimal parameters that can be employed for a wide range of systems in quantum Monte Carlo calculations. Chapter 3 introduces a simple, yet general, method for constructing basis sets appropriate for molecular electronic structure calculations. This method is employed to create a standard Gauss-Slater basis set for each of the elements from hydrogen to argon.

1.5.3 Overview of Chapter 4

Chapter 4, traverses a different path from Chapters 2 and 3. Instead of focusing on improving a particular aspect of the trial wavefunction, this project explores the effect of different aspects of the trial wavefunction on the accuracy of quantum Monte Carlo. By systematically testing the effect of the basis size, orbital quality, and determinant expansion quality, this work offers guidance to quantum Monte Carlo practitioners for achieving results to within chemical accuracy of experiment.

1.5.4 Overview of Chapter 5

In Chapter 5, semistochastic projection, a hybrid of deterministic projection (exact diagonalization) and stochastic projection (quantum Monte Carlo), is introduced for finding the dominant eigenvalue and eigenvector of a matrix. Before this work, projection methods have applied the projector deterministically or stochastically via Monte Carlo. Due to the superior computational efficiency of deterministic projection, stochastic projection is reserved for spaces

with more states than can be stored in memory. However, even if the entire space is too large for deterministic projection, a subset of the space can be handled with deterministic projection; and the rest of the space can be handled with stochastic projection. This is the essence of semistochastic projection.

After introducing the algorithm, the method is applied to determining the ground state energy of the Hamiltonian in a discrete basis. This application of semistochastic projection, dubbed semistochastic quantum Monte Carlo, is shown to be orders of magnitude more efficient than stochastic quantum Monte Carlo.

CHAPTER 2

COMPACT AND FLEXIBLE BASIS FUNCTIONS FOR QUANTUM MONTE CARLO CALCULATIONS

The text of this chapter is a reproduction of a paper written on the same subject in 2010 [7]. The reference to the paper is J. Chem. Phys. 132, 094109 (2010). The abbreviations used in this chapter are given in Table 2.1.

Table 2.1: Abbreviations for Chapter 2. Descriptions are provided for non-standard abbreviations.

Abbreviation	Description
QC	Quantum Chemistry
QMC	Quantum Monte Carlo
HF	Hartree-Fock
GS	Gauss-Slater: Basis function that behaves like a Gaussian at short distances and a Slater at long distances
VMC	Variational Monte Carlo
DMC	Diffusion Monte Carlo
BFD	Burkatzki, Filippi and Dolg: Gaussian basis sets and non-divergent pseudopotentials constructed for QMC
CSF	Configuration state function
CAS	Complete Active Space
CCSD(T)	Coupled Cluster with single and double excitations and perturbative triple excitations
RMS	Root-mean-square

2.1 Introduction

In traditional quantum chemistry (QC) calculations, molecular orbitals are often expanded in a combination of contracted Gaussian basis functions and primitive Gaussian basis functions. For each occupied orbital, a contracted function is constructed to reproduce the corresponding atomic orbital from an effectively single-electron theory such as Hartree-Fock (HF) [8, 9], or the natural orbital from a post-HF method [10, 11].

While a single primitive Gaussian has incorrect long-range asymptotic behavior, a contracted basis function can reproduce the correct asymptotics over a reasonable range. However, even contracted functions are unable to produce the correct electron-nucleus cusps [12] since they have zero gradient at the origin. Despite these shortcomings, Gaussians are used in QC calculations because they permit analytical evaluation of the two-electron integrals [13].

In contrast to traditional QC methods, quantum Monte Carlo (QMC) calculations [4] enjoy greater wavefunction flexibility by using Monte Carlo integration to evaluate matrix elements. In particular, bases need not be restricted to Gaussians. For calculations employing a potential that diverges at the nucleus, Slater basis functions can exactly reproduce the correct electron-nucleus cusp and long-range asymptotic behavior of the orbitals. In fact, for all-electron QMC calculations, highly accurate results have been obtained by employing compact basis sets consisting of Slater functions with optimized exponents [14, 15].

Conversely, the basis sets used for non-divergent pseudopotential calculations in QMC have deviated little from typical QC basis sets. For these pseudopotentials, orbitals have no electron-nucleus cusp. In this case, Gaussian basis

functions are more appropriate than Slater functions at small electron-nuclear distances but still have incorrect long-range asymptotics.

Contracted and primitive Gaussian functions are frequently splined on a radial grid for QMC. Splining contracted Gaussians presents a definite computational advantage since evaluating polynomials is much cheaper than evaluating a linear combination of Gaussians. In contrast, splining primitive Gaussians provides minimal benefit at best.

We propose two ideas for improving basis sets for pseudopotential calculations in QMC. First, primitive basis function exponents are optimized for each system. This provides greater accuracy with a compact basis for a wide range of chemical environments and excitation levels. To facilitate optimization, the primitive basis functions remain analytic while the contracted functions are splined.

Second, we propose a novel form of primitive basis function appropriate for calculations involving non-diverging pseudopotentials. These primitives, which we call Gauss-Slater (GS) functions, have the short-range behavior of a Gaussian function and the long-range behavior of a Slater function.

The utility of our improvements is demonstrated by calculations for carbon, the lowest lying excited states of carbon with ${}^5S^o$, ${}^3P^o$, ${}^1D^o$, ${}^3F^o$ symmetries, carbon dimer, and naphthalene.

This paper is organized as follows. In Section 2.2, the form and properties of Gauss-Slater functions are introduced. In Section 2.3, results of our calculations are discussed. In Section 2.4, concluding remarks are provided. In Appendix B, technical details are discussed.

2.2 Gauss-Slater Basis Functions

We define Gauss-Slater (GS) functions as

$$\varphi^{\zeta}_{nlm}(r,\theta,\phi) = N^{\zeta}_{n}\; r^{n-1}e^{-\frac{(\zeta r)^2}{1+\zeta r}}\; Z^{m}_{l}(\theta,\phi), \tag{2.1}$$

where r, θ, ϕ are the standard spherical coordinates, n is the principal quantum number, l is the azimuthal quantum number, m is the magnetic quantum number, N^{ζ}_{n} is the normalization factor, and $Z^{m}_{l}(\theta,\phi)$ is a real spherical harmonic.

Notice that for $r \ll 1$ the GS behaves like a Gaussian:

$$\varphi^{\zeta}_{nlm}(r,\theta,\phi) \cong N^{\zeta}_{n}\; r^{n-1}e^{-(\zeta r)^2}\; Z^{m}_{l}(\theta,\phi), \tag{2.2}$$

and for $r \gg 1$ the GS behaves like a Slater:

$$\varphi^{\zeta}_{nlm}(r,\theta,\phi) \cong N^{\zeta}_{n}\; r^{n-1}e^{-\zeta r}\; Z^{m}_{l}(\theta,\phi). \tag{2.3}$$

The GS drift velocity and local energy are well behaved at long distances, while for Gaussians these quantities diverge.

Unlike Gaussians and Slaters, normalization of GSs has no closed form expression. Nevertheless, normalizing an arbitrary GS is trivial with the following scaling relation (see Appendix B) between N^{ζ}_{n} and N^{1}_{n},

$$N^{\zeta}_{n} = \zeta^{n+1/2}\; N^{1}_{n}. \tag{2.4}$$

Since GSs are not analytically integrable, the exponential part must be expanded in Gaussians for use in quantum chemistry programs that employ analytic integrals for evaluating the matrix elements. This expansion is

$$N^{\zeta}_{n}\; e^{-\frac{(\zeta r)^2}{1+\zeta r}} = \sum_{i} c^{\zeta}_{i}\; \sqrt{\frac{2(2\alpha^{\zeta}_{i})^{n+\frac{1}{2}}}{\Gamma(n+\frac{1}{2})}}\; e^{-\alpha^{\zeta}_{i} r^2}, \tag{2.5}$$

where c_i^ζ is the i^{th} expansion coefficient, and α_i^ζ is the i^{th} Gaussian exponent. Additionally, the following scaling relations (see Appendix B) hold for the expansion coefficients and Gaussian exponents:

$$\alpha_i^\zeta = \zeta^2 \alpha_i^1 \tag{2.6}$$

$$c_i^\zeta = c_i^1. \tag{2.7}$$

Once the Gaussian expansions are found for unit exponents, expansions of arbitrary GSs follow immediately from the scaling relations.

2.3 Results

For all applications discussed in this paper, variational Monte Carlo (VMC) and diffusion Monte Carlo (DMC) [16] calculations are performed with the CHAMP QMC code [17] and employ the pseudopotentials and accompanying basis sets of Burkatzki, Filippi and Dolg (BFD) [18]. We choose these pseudopotentials and basis functions since they were constructed for use in QMC and have proved to be quite accurate.

The wavefunction is of the standard Slater-Jastrow form. All wavefunction parameters including Jastrow parameters, Configuration State Function (CSF) coefficients (where applicable), orbital coefficients, and primitive exponents (where applicable) are optimized via the linear method [19, 20, 21]. Optimization is performed on a linear combination of the energy and variance of the local energy with weights 0.95 and 0.05, respectively. Optimizing just the energy yields slightly lower energies and somewhat higher variances.

For each system considered, calculations are performed with three different

basis sets: (1) the BFD basis, (2) fixed contracted functions and analytical Gaussian primitives with optimized exponents, and (3) fixed contracted functions and analytical Gauss-Slater primitives with optimized exponents. We refer to these basis sets as BFD, G, and GS, respectively.

These three cases allow us to evaluate the improvements our two methods provide to the current basis sets used in QMC. First, if both the G and GS basis sets significantly outperform the BFD basis, then the utility of reoptimizing the basis exponents within QMC will be established. Second, the utility of the GS basis depends on its performance relative to the G basis.

2.3.1 Ground State Carbon Atom

For the carbon atom ground state, ^3P, we consider a complete active space (CAS) wavefunction with an active space generated by distributing the four valence electrons among the thirteen orbitals of the $n = 2$ and $n = 3$ shells. Denoted by CAS(4,13), this wavefunction consists of 83 CSFs comprised of 422 determinants.

In general, a single Slater determinant will not be a CSF when a certain number of electrons have been excited relative to the ground state HF Slater determinant. However, a CSF can be produced from an arbitrary Slater determinant by applying projection operators for angular momentum $\hat{L}$ and spin $\hat{S}$. Since states with the same L and S but different L_z and S_z are degenerate, we are free to choose convenient L_z and S_z states. We choose $L_z = 0$ to make the wavefunctions real to within a position independent phase, and we choose $S_z = S$ to yield the minimum number of determinants in the CSF. Since the carbon ground state

has $L = 1$, $S = 1$, the projection operators are of the form

$$\hat{P}_L = \prod_{L' \neq 1} \left[\hat{L} - L'(L'+1) \right] \tag{2.8}$$

$$\hat{P}_S = \prod_{S' \neq 1} \left[\hat{S} - S'(S'+1) \right], \tag{2.9}$$

where the product over all possible angular momentum and spin values omits the desired $L = 1$ and $S = 1$ values.

Carbon atom VMC results for each basis set are shown in Table 2.2. Included for comparison, coupled cluster calculations with single and double excitations and perturbative triple excitations (CCSD(T)) values for the BFD basis [18] exhibit much larger dependence on basis size than QMC results.

Both the G and GS basis sets outperform the BFD basis set. The $2z$ G basis exhibits a modest gain of 0.3 mH in energy compared to the corresponding BFD basis. The $2z$ GS basis exhibits larger gains of 1 mH in energy and 28 mH in σ, the root-mean-square (RMS) fluctuations of the local energy. The $3z$ GS basis yields identical results, within statistical error, to the $5z$ BFD basis.

Carbon atom DMC results for each basis set are shown in Table 2.3. These calculations were performed with a time step of $\tau = 0.01$ H^{-1} which leads to negligible time step error for these high quality wavefunctions. DMC depends less on basis size than VMC, as is immediately apparent from the data. Nevertheless, both the G and GS basis sets outperform the BFD basis set. The $3z$ GS basis yields identical results, within statistical error, to the $5z$ BFD basis, and an energy 0.1 mH lower than the $3z$ BFD basis.

Both VMC and DMC results indicate that reoptimizing primitive basis function exponents provides improvements which can be significant for the GS ba-

sis. In large systems, the ability to use a $3z$ basis in place of a $4z$ or $5z$ basis determines whether a calculation can be performed.

Table 2.2: VMC energy and RMS fluctuations of the local energy, σ, in Hartrees for CAS(4,13) ground state of carbon using BFD, G, and GS basis functions. Statistical errors on the last digit are shown in parentheses. For each n, the nz basis consists of n S functions, n P functions, and $n-1$ D functions. CCSD(T) values for the BFD basis are included for comparison [18].

Type	Size	Energy (H)	σ (H)
BFD	$2z$	-5.43161(3)	0.1395(6)
	$3z$	-5.43306(2)	0.099(3)
	$4z$	-5.43332(2)	0.0904(2)
	$5z$	-5.43341(2)	0.0905(4)
G	$2z$	-5.43196(3)	0.138(2)
	$3z$	-5.43324(2)	0.0989(5)
GS	$2z$	-5.43264(2)	0.1114(4)
	$3z$	-5.43344(2)	0.0898(2)
CCSD(T)	$2z$	-5.409230	N/A
	$3z$	-5.427351	N/A
	$4z$	-5.431486	N/A
	$5z$	-5.432494	N/A

2.3.2 Excited States of Carbon

We consider the lowest lying excited states of carbon with ${}^5S^o$, ${}^3P^o$, ${}^1D^o$, and ${}^3F^o$ symmetries. These states have configurations $2s^1 2p^3$, $2s^2 2p^1 3s^1$, $2s^2 2p^1 3d^1$,

Table 2.3: DMC energy in Hartrees for CAS(4,13) ground state of carbon using BFD, G, and GS basis functions. Statistical errors on the last digit are shown in parentheses. For each n, the nz basis consists of n S functions, n P functions, and $n-1$ D functions. Calculations were performed with a time step of $\tau = 0.01$ H^{-1} which leads to a negligible time step error for these high quality wavefunctions.

Type	Size	Energy (H)
BFD	$2z$	-5.43314(2)
	$3z$	-5.43395(2)
	$4z$	-5.43404(1)
	$5z$	-5.43407(1)
G	$2z$	-5.43342(2)
	$3z$	-5.43400(2)
GS	$2z$	-5.43356(2)
	$3z$	-5.43407(1)

and $2s^2 2p^1 3d^1$, respectively. The $^3P^o$, $^1D^o$, and $^3F^o$ states have much higher energy than the ground state and $^5S^o$ excited state.

The dominant CSF for each of these three states occupies orbitals that are unoccupied in the HF ground state. For fair comparison, the BFD basis therefore must be augmented. The diffuse functions of the aug-cc-pVnZ basis sets [11, 22, 23] are used for this purpose. The BFD nz basis then becomes an $(n+1)z$ basis.

Obtained by application of the projection operators discussed in Section 2.3.1, the dominant CSF for each of the four excited states has one, one, four, and three Slater determinants, respectively.

VMC results for energies and σ of each system are shown in Figures 2.1 and

2.2. In all cases, the reoptimized exponents provide significant gains in both energy and σ. Results for the three higher lying states demonstrate that reoptimized exponents are essential for describing states containing orbitals unoccupied in the HF ground state. For these systems, $2z$ results using the G and GS basis sets are substantially better than $5z$ results for the augmented BFD basis set. In the most extreme case of $^3F^o$, the $2z$ GS basis results in 30 mH lower energy and 110 mH lower σ than the $5z$ BFD basis.

The importance of the reoptimized exponents is evident for the excited states of carbon. However, benefits of the GS basis relative to the G basis are never more than several tenths of a mH. On the scale of the plots in Figures 2.1 and 2.2, many G and GS basis results coincide.

2.3.3 Carbon Dimer

Single determinant calculations were performed for the carbon dimer with initial wavefunctions generated from the QC code GAMESS [24].

VMC results for each basis set are shown in Table 2.4. The G and GS basis sets outperform the BFD basis set. In particular, the $2z$ G basis attains a 0.6 mH lower energy than the corresponding BFD basis, and the GS basis yields a 3.2 mH lower energy than the BFD basis. The $3z$ GS basis yields an energy within 0.3 mH of and a σ identical to the $5z$ BFD basis results.

DMC results for each basis set are shown in Table 2.5. These calculations were performed with a time step of $\tau = 0.005$ H^{-1} which leads to negligible time step error. The $2z$ G and GS basis sets significantly outperform the cor-

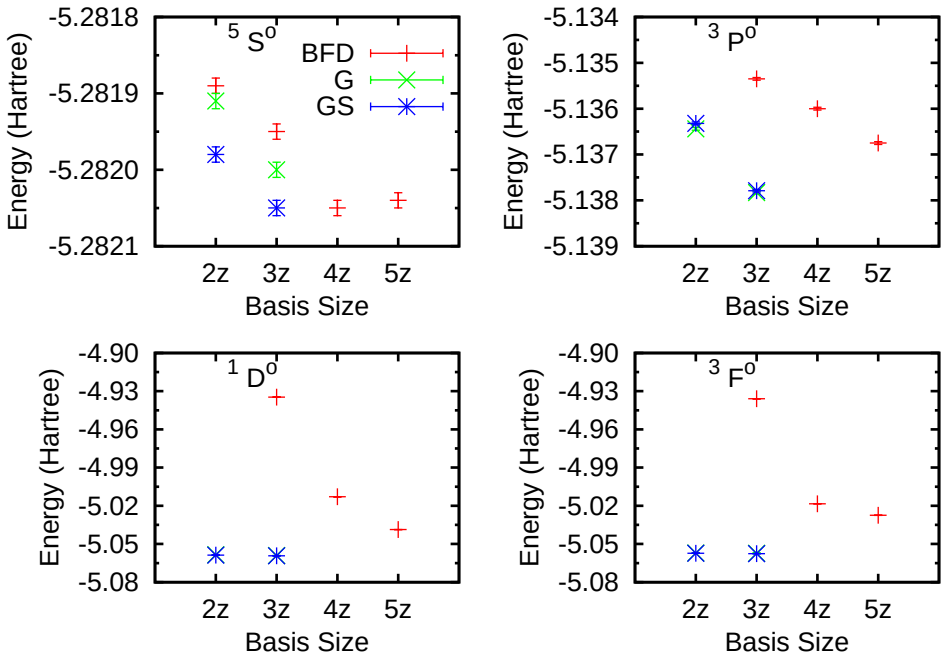

Figure 2.1: VMC energies in Hartrees for lowest lying excited states of carbon with ${}^5S^o$, ${}^3P^o$, ${}^1D^o$, ${}^3F^o$ symmetries. For each n, the nz basis consists of n S functions, n P functions, and $n-1$ D functions (where applicable). For ${}^3P^o$, ${}^1D^o$, ${}^3F^o$ calculations, the BFD basis is augmented with diffuse functions of the aug-cc-pVnZ basis sets [11, 22, 23]. For G and GS basis sets, only $2z$ and $3z$ calculations were performed. In many cases, results for G and GS bases are indistinguishable on this scale.

responding BFD basis set. The $2z$ GS basis yields a result that is essentially converged with respect to basis size.

Both VMC and DMC results indicate that reoptimizing primitive basis function exponents provides improvements which can be significant for the GS basis.

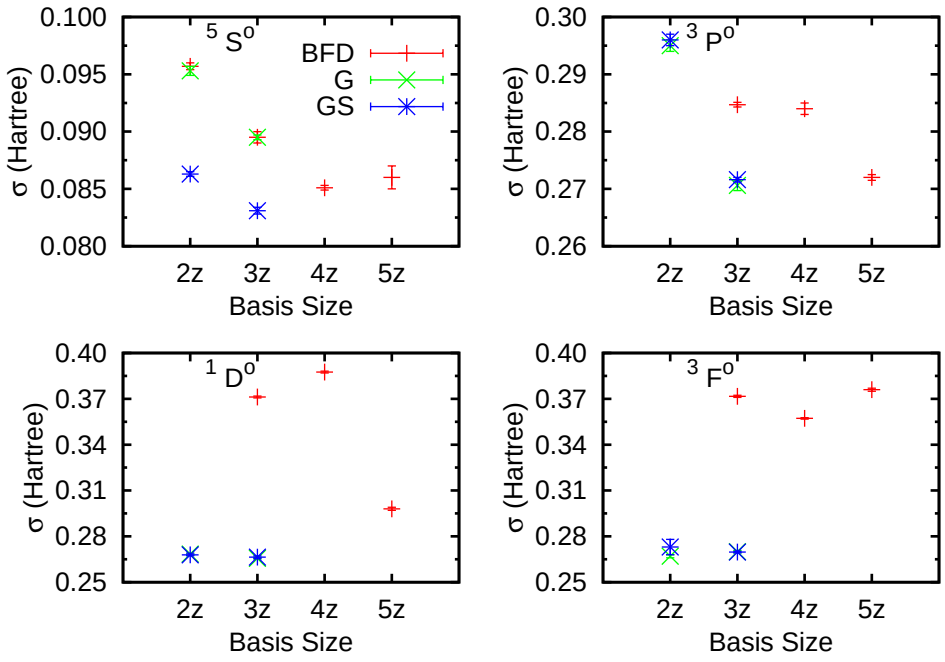

Figure 2.2: RMS fluctuations of VMC local energies. See Figure 2.1 for notation and details.

2.3.4 Naphthalene

Single determinant calculations were performed for naphthalene, $C_{10}H_8$, with initial wavefunctions generated from the QC code GAMESS [24].

Calculations were performed only for the $2z$ basis. The intention of this section is not to produce an energy converged with respect to basis size, but to demonstrate that the utility of reoptimizing primitive basis functions, and GS primitives in particular, extends to large systems. VMC and DMC results for each basis set are shown in Table 2.6. The DMC calculations were performed with a time step of $\tau = 0.01$ H^{-1}.

At both the VMC and DMC level, wavefunction quality increases by reoptimizing the primitive Gaussian exponents. The GS basis provides further im-

Table 2.4: VMC energy and σ in Hartrees for single determinant carbon dimer ground state. Statistical errors on the last digit are shown in parentheses. For each n, the nz basis consists of n S functions, n P functions, $n-1$ D functions, $n-2$ F functions, and $n-3$ G functions.

Type	Size	Energy (H)	σ (H)
BFD	$2z$	-11.02644(4)	0.4343(9)
	$3z$	-11.03003(4)	0.4172(4)
	$4z$	-11.03094(4)	0.4127(7)
	$5z$	-11.03095(4)	0.4113(6)
G	$2z$	-11.02707(4)	0.4288(7)
	$3z$	-11.03030(4)	0.4183(6)
GS	$2z$	-11.02968(4)	0.4191(6)
	$3z$	-11.03065(4)	0.4109(6)

provement. In particular, even at the DMC level, the $2z$ G basis attains a 4 mH lower energy than the corresponding BFD basis, and the GS basis yields a 15 mH lower energy than the BFD basis. This is significant since DMC is less sensitive to basis set choice than VMC.

2.4 Conclusion

Quantum Monte Carlo (QMC) methods have the advantage over standard quantum chemistry (QC) methods of rapid convergence with increasing basis size. Basis-size dependence for pseudopotential calculations in QMC is further reduced by two basis set improvements introduced in this work. Calculations

Table 2.5: DMC energy in Hartrees for ground state of carbon dimer using BFD, G, and GS basis functions. Statistical errors on the last digit are shown in parentheses. For each n, the nz basis consists of n S functions, n P functions, $n-1$ D functions, $n-2$ F functions, and $n-3$ G functions. Calculations were performed with a time step of $\tau = 0.005$ H^{-1} which leads to negligible time step error.

Type	Size	Energy (H)
BFD	$2z$	-11.05561(3)
	$3z$	-11.05719(4)
	$4z$	-11.05728(4)
	$5z$	-11.05723(4)
G	$2z$	-11.05632(3)
	$3z$	-11.05717(3)
GS	$2z$	-11.05702(4)
	$3z$	-11.05719(3)

Table 2.6: VMC energy and σ, and DMC energy in Hartrees for single determinant ground state naphthalene, $C_{10}H_8$. Statistical errors on the last digit are shown in parentheses. DMC calculations were performed with a time step of $\tau = 0.01$ H^{-1}. For carbon, the $2z$ basis includes 2 S function, 2 P function, 1 D function. For hydrogen, the $2z$ basis includes 2 S functions and 1 P function.

Type	Size	VMC Energy (H)	VMC σ (H)	DMC Energy (H)
BFD	$2z$	-61.5193(5)	0.980(1)	-61.6479(5)
G	$2z$	-61.5273(4)	0.938(1)	-61.6518(5)
GS	$2z$	-61.5438(4)	0.927(2)	-61.6634(5)

for the ground state of carbon, the lowest lying $^5S^o$, $^3P^o$, $^1D^o$, $^3F^o$ excited states of carbon, carbon dimer, and naphthalene demonstrate the utility of our contribution.

First, we reoptimized the primitive basis function exponents for each system because the exponents of standard QC and QMC basis sets, such as the Burkatzki, Filippi and Dolg (BFD) basis, represent a compromise. These standard exponents are designed to yield good energies for some range of chemical environments and excitation levels, but they cannot be optimal for all systems. We have shown that reoptimizing primitive basis function exponents for each system yields significant improvements in the energy and fluctuations of the local energy, σ. The most pronounced benefits were observed in higher-lying excited state calculations. In the most extreme case of $^3F^o$ at the variational Monte Carlo (VMC) level, the $2z$ mixed basis was 30 mH lower in energy and 110 mH lower in σ than the $5z$ numerical basis. Although not discussed in this paper, we have found that reoptimization of standard Slater basis exponents used in all-electron calculations also provides considerable improvements in energy and σ.

Second, we introduced Gauss-Slater (GS) basis functions for non-divergent pseudopotential calculations. GS functions behave like Gaussians at short distances and Slaters at long distances. In all systems considered, results obtained using a mixed basis comprised of contracted and primitive basis functions improved when optimized Gaussian primitives were replaced by optimized GS primitives. Importantly, for carbon dimer at the DMC level the $2z$ GS total energies are nearly converged with respect to basis size.

A $3z$ mixed basis with optimized GSs for carbon atom or carbon dimer pro-

duces results comparable to the $5z$ BFD basis. Since the number of orbital coefficients to be optimized scales quadratically with basis size, the use of a more compact basis allows larger problems to be attacked in QMC.

CHAPTER 3

BASIS SET CONSTRUCTION FOR MOLECULAR ELECTRONIC STRUCTURE THEORY: NATURAL ORBITAL AND GAUSS-SLATER BASIS FOR SMOOTH PSEUDOPOTENTIALS

The text of this chapter is a reproduction of a paper written on the same subject in 2011 [25]. The reference to the paper is J. Chem. Phys. 134, 064104 (2011). The abbreviations used in this chapter are given in Table 3.1.

3.1 Introduction

In quantum chemistry (QC) calculations, molecular orbitals are traditionally expanded in a combination of primitive Gaussian basis functions and linear combinations of Gaussian primitives called contracted basis functions [10]. These basis sets cannot express the correct molecular orbital asymptotic behavior but are used in QC calculations to permit analytic evaluation of the two-electron integrals [13].

Analytic integral evaluation significantly limits flexibility in basis set choice but is essential for computational efficiency in QC calculations. However, in practice, other basis function forms can be considered since an arbitrary function can be expanded in Gaussians. Of course, the fidelity of this representation is limited. An expansion in a finite number of Gaussians cannot reproduce the exponential decay of the wavefunction at large distances or the Kato cusp conditions [12] at nuclei, but it can mimic these features over a finite range.

Quantum Monte Carlo (QMC) calculations [4] offer greater freedom in choice

Table 3.1: Abbreviations for Chapter 3. Descriptions are provided for non-standard abbreviations.

Abbreviation	Description
QC	Quantum Chemistry
QMC	Quantum Monte Carlo
GS	Gauss-Slater: Basis function that behaves like a Gaussian at short distances and a Slater at long distances
NOs	Orbitals obtained by diagonalizing the density matrix
HF	Hartree-Fock
MCSCF	Multiconfigurational self-consistent field
CCSD	Coupled Cluster with single and double excitations
BFD	Burkatzki, Filippi and Dolg: Gaussian basis sets and non-divergent pseudopotentials constructed for QMC
B3LYP	Becke three-parameter hybrid density functional
nz	$n-$zeta basis
CAS	Complete Active Space
ANO-GS	Basis with contractions that are atomic natural orbitals and primitives that are Gauss-Slaters
DMC	Diffusion Monte Carlo
CSF	Configuration state function

of basis functions because matrix elements are evaluated using Monte Carlo integration. Consequently, the correct short- and long-distance asymptotics can be satisfied exactly. For systems with a divergent nuclear potential, Slater basis functions can exactly reproduce the correct electron-nucleus cusp and long-range asymptotic behavior of the orbitals. For calculations on systems with a potential that is finite at the nucleus and has a Coulomb tail, Gauss-Slater (GS) primitives [7] are the appropriate choice since they introduce no cusp at the origin and reproduce the exponential long-range asymptotic behavior of the orbitals.

Despite shortcomings, traditional QC basis sets have yielded good results. The natural orbitals (NOs) from a post Hartree-Fock (HF) method are a particularly successful form of contracted function [26, 27, 28, 29]. The simplest NO construction involves diagonalizing the one-particle density matrix from a ground state atomic calculation [26]. This construction is unbalanced due to obvious bias favoring the atom. More complicated constructions involve diagonalizing the average one-particle density matrix of several systems: atomic ground and excited states, ions, diatomic molecules, and atoms in an external electric field [27, 28, 29]. These constructions produce excellent results, but they are complex.

A simple but general method for constructing basis sets for molecular electronic structure calculations is proposed and tested here. The bases are combinations of the NOs obtained from diagonalizing the one-particle density matrix from an atomic multiconfigurational self-consistent field (MCSCF) calculation and primitive functions appropriate for the potential in the system. The primitives are optimized for the homonuclear dimer in coupled cluster calculations

with single and double excitations (CCSD), with the intention of producing a balanced basis set. Importantly, optimal exponents for the primitive functions are shown to depend weakly on the level of theory used in the optimization. Additionally, results show that coupling is weak between primitive functions of different angular momenta. This enables efficient determination of optimal exponents.

The utility of the above construction is demonstrated for the elements hydrogen through argon with the non-divergent pseudopotentials of Burkatzki, Filippi, and Dolg (BFD) [18]. Since these pseudopotentials are finite at the nuclei and have a Coulomb tail, the GS functions are the appropriate primitives. These pseudopotentials are chosen for demonstrated accuracy in all cases tested and because they are accompanied by a basis set. The BFD basis [18] serves as a metric for testing the new basis. The benefits of our bases extend to all electronic structure methods tested, including CCSD, HF, the Becke three-parameter hybrid density functional (B3LYP) [30], and QMC.

The main area of interest for the authors is QMC. Since QMC results depend less on basis set than traditional QC methods [7], only double-zeta ($2z$) and triple-zeta ($3z$) bases are presented.

This paper is organized as follows. Basis function form and properties are demonstrated in Sec. 3.2. Results for calculations with the new bases are discussed in Sec. 3.3. Concluding remarks are provided in Sec. 3.4. Supplementary material is provided on EPAPS [31] and in Appendix C.

3.2 Basis Set

The number of basis functions for each angular momentum follows the correlation consistent polarized basis set prescription of Dunning [10]. $2z$ and $3z$ bases appropriate for the BFD pseudopotentials are generated for the elements hydrogen through argon. Since the BFD pseudopotential removes no core for hydrogen and helium, the $2z$ basis for these elements consists of two S functions and one P function, while the $3z$ basis consists of three S functions, two P functions, and one D function. Since the BFD pseudopotential removes a helium core for the first row atoms and a neon core for the second row atoms, the remaining elements lithium through argon have the same number of basis functions. In particular, the $2z$ basis consists of two S functions, two P functions, and one D function, while the $3z$ basis consists of three S functions, three P functions, two D functions, and one F function.

The bases consist of a combination of contracted and primitive functions. Since the BFD pseudopotentials are finite at the origin and have a Coulomb tail, the GS functions are the appropriate primitives. With the exception of the elements in Group $1A$ of the periodic table (i.e. H, Li, and Na), the basis for each element includes a single S contraction and a single P contraction combined with an appropriate number of GS primitives. Only two contractions are employed to reduce the computational cost of using this basis in QC calculations. Since elements in Group $1A$ of the periodic table have only one electron for the BFD pseudopotentials, a single S orbital is the ground state wavefunction, and this can be obtained exactly in HF. Thus, the basis for each element in Group $1A$ includes a single S contraction, no P contractions, and an appropriate number of GS primitives.

3.2.1 Contracted Functions

A contracted basis function is a linear combination of Gaussian primitives:

$$\varphi_{nlm}(r,\theta,\phi) = \sum_i c_i \sqrt{\frac{2(2\alpha_i)^{n+\frac{1}{2}}}{\Gamma(n+\frac{1}{2})}}\, r^{n-1} e^{-\alpha_i r^2} Z_l^m(\theta,\phi), \tag{3.1}$$

where r, θ, ϕ are the standard spherical coordinates, n is the principal quantum number, l is the azimuthal quantum number, m is the magnetic quantum number, $Z_l^m(\theta,\phi)$ is a real spherical harmonic, c_i is the i^{th} expansion coefficient, and α_i is the i^{th} Gaussian exponent. In practice, the restriction $n = l + 1$ applies.

The exponents of the primitive functions that form the contracted basis functions are determined as follows. For each angular momentum for which a contraction is desired, an uncontracted basis consisting of nine even-tempered primitive Gaussians is generated. For each set of uncontracted Gaussians, the minimum exponent and even-tempering coefficient are varied to minimize the CCSD energy of the atom using a Python wrapper around GAMESS [24].

An assumption of weak coupling between the different angular momenta underlies the optimization procedure. Consequently, the uncontracted basis for each angular momentum is optimized separately. This optimization is performed by calculating the CCSD energy on an initially coarse grid composed of different minimum exponents and even-tempering coefficients. Once regions of low CCSD energy are identified, a finer grid is used to obtain the final minimum exponent and even-tempering coefficient. In addition to the assumption of weak coupling, two other properties of the problem make this global optimization possible with modest computer resources; low dimensionality of search space and efficiency of atomic CCSD calculations.

Next, an atomic MCSCF calculation in a complete active space (CAS) with

the optimized uncontracted basis is performed in GAMESS. For these calculations, all electrons not removed by the pseudopotential are allowed to excite. For helium, the active space consists of the orbitals from the $n = 1$ and $n = 2$ shells. For beryllium through neon, the active space includes the orbitals from the $n = 2$ and $n = 3$ shells. For magnesium through argon, the active space is composed of the orbitals from the $n = 3$ and $n = 4$ shells, with the exception of the 4D and 4F orbitals. A subset of the natural orbitals from the MCSCF calculations are used as the contracted functions of our basis.

All atomic calculations are performed in D_{2h} symmetry since GAMESS does not permit imposition of full rotational symmetry. Hence, different components of the same atomic subshell are not necessarily equivalent. Additionally, mixing may occur among orbitals of different angular momenta. For instance, there is mixing of S orbitals with both $D_{3z^2-r^2}$ and $D_{x^2-y^2}$ orbitals. This anisotropy can be removed by averaging the different components of a particular subshell and zeroing out the off-diagonal blocks of the one-particle density matrix [27].

A simpler approach taken in this work is found to produce results of similar quality. For each angular momentum for which a contraction is desired, the NO with that angular momentum which has the largest occupation number is chosen. Additionally, NO elements which do not correspond to the dominant character of the orbital are zeroed out. For instance, an NO with large coefficients on the S basis functions and small coefficients on the D basis functions is considered to be dominated by S character, so the D coefficients are zeroed out. Finally, the NOs are normalized. The NOs selected in this procedure generate the contracted functions for the basis set. The expansions of the contractions are given in the supplementary material [31] and Appendix C.

3.2.2 Gauss-Slater Primitives

GS functions [7] are defined as

$$\varphi^{\zeta}_{nlm}(r,\theta,\phi) = N^{\zeta}_{n}\; r^{n-1}e^{-\frac{(\zeta r)^2}{1+\zeta r}}\; Z^{m}_{l}(\theta,\phi), \tag{3.2}$$

where ζ is the GS exponent and N^{ζ}_{n} is the normalization factor. The restriction $n \geq l+1$ is imposed for GS functions. For $r \ll 1$, the GS behaves like a Gaussian:

$$\varphi^{\zeta}_{nlm}(r,\theta,\phi) \cong N^{\zeta}_{n}\; r^{n-1}e^{-(\zeta r)^2}\; Z^{m}_{l}(\theta,\phi), \tag{3.3}$$

and for $r \gg 1$, the GS behaves like a Slater:

$$\varphi^{\zeta}_{nlm}(r,\theta,\phi) \cong N^{\zeta}_{n}\; r^{n-1}e^{-\zeta r}\; Z^{m}_{l}(\theta,\phi). \tag{3.4}$$

Consequently, GS functions introduce no cusp at the origin and can reproduce correct long-range asymptotic behavior of the orbitals.

Unlike Gaussians and Slaters, normalization of GSs has no closed form expression. Nevertheless, normalizing an arbitrary GS is trivial with the following scaling relation between N^{ζ}_{n} and N^{1}_{n}:

$$N^{\zeta}_{n} = \zeta^{n+1/2}\; N^{1}_{n}. \tag{3.5}$$

Values for N^{1}_{n} are given in the supplementary material [31] and Appendix C.

Since GSs are not analytically integrable, the radial part must be expanded in Gaussians for use in QC programs that evaluate matrix elements analytically. The expansion is

$$\varphi^{\zeta}_{nlm}(r,\theta,\phi) = \sum_{i} c^{\zeta}_{i}\; \sqrt{\frac{2(2\alpha^{\zeta}_{i})^{l+\frac{3}{2}}}{\Gamma(l+\frac{3}{2})}}\; r^{l}e^{-\alpha^{\zeta}_{i}r^2}\, Z^{m}_{l}(\theta,\phi), \tag{3.6}$$

where c^{ζ}_{i} is the i^{th} expansion coefficient and α^{ζ}_{i} is the i^{th} Gaussian exponent. Notice that the expansion permits the case for which $n \neq l+1$ for the GS function.

Additionally, the following scaling relations hold for the expansion coefficients and Gaussian exponents:

$$\alpha_i^\zeta = \zeta^2 \alpha_i^1 \tag{3.7}$$

$$c_i^\zeta = c_i^1. \tag{3.8}$$

Once the Gaussian expansions are found for unit exponents, expansions of arbitrary GSs follow immediately from the scaling relations. For QC calculations in this paper, GSs are expanded in six Gaussians. However, if the purpose of the initial QC calculation is to generate crude starting orbitals for QMC calculations in which orbital optimization is performed, it is only necessary to expand GS primitives in a single Gaussian. In this case, the cost of QC calculations is the same for Gaussian and GS primitives. The expansions of GS functions with unit exponent in both one and six Gaussians are given in the supplementary material [31] and Appendix C.

As mentioned above, the restriction $n \geq l + 1$ is imposed for GS functions, instead of the more familiar $n = l + 1$ restriction imposed for Gaussian primitives. This motivates construction of two types of bases. In the first, ANO-GS, the restriction $n = l + 1$ is enforced. In the second, ANO-GSn, for each l there can be at most a single GS primitive with a particular n. For each additional primitive with a particular l, n must be incremented.

For example, consider lithium. The $2z$ ANO-GS basis has one S contraction, one GS-1S function, two GS-2P functions, and one GS-3D function. On the other hand, the $2z$ ANO-GSn basis has one S contraction, one GS-1S function, one GS-2P function, one GS-3P function, and one GS-3D function.

A caveat to the above definition of the ANO-GSn basis is that GS-2S func-

tions are not permitted since a single GS-2S function will introduce an undesired cusp in the wavefunction. Additionally, the $2z$ ANO-GS and ANO-GSn basis sets are identical for all elements except lithium and sodium. When the $2z$ ANO-GS and ANO-GSn basis sets are identical, the basis sets are referred to as a $2z$ ANO-GS/GSn basis. For both lithium and sodium, the basis sets differ because these systems have no P contractions and instead have a second P primitive for the $2z$ basis. This primitive is a GS-2P for the ANO-GS basis and a GS-3P for the ANO-GSn basis. Additionally, weak coupling between functions of different angular momentum causes the GS-1S and GS-3D functions in the ANO-GS bases for lithium and sodium to differ from their counterparts in the ANO-GSn bases. However, the optimal exponents differ by less than 0.01.

Optimal exponent selection for the GS primitives is discussed now. Instead of optimizing exponents for the atom as was done to generate the contractions, optimization of the GS exponents is performed for the homonuclear diatomic molecule at experimental bond length [32, 33, 34, 35, 36, 37, 38, 39, 40]. This advantageously produces a balanced basis set.

Weak coupling between GS functions of different angular momenta is assumed, so the initial optimization for each angular momentum is performed separately. This assumption is validated in Figure 3.1, which contains plots of the CCSD energy for Si_2 while varying individual GS exponents in the $2z$ ANO-GS/GSn basis. Both the curve shape and exponent value which minimizes the energy vary little with fixed exponent value, signifying weak coupling between GS functions of different angular momentum.

The optimization is performed at the CCSD level of theory using a Python wrapper around GAMESS. For each angular momentum, an energy landscape

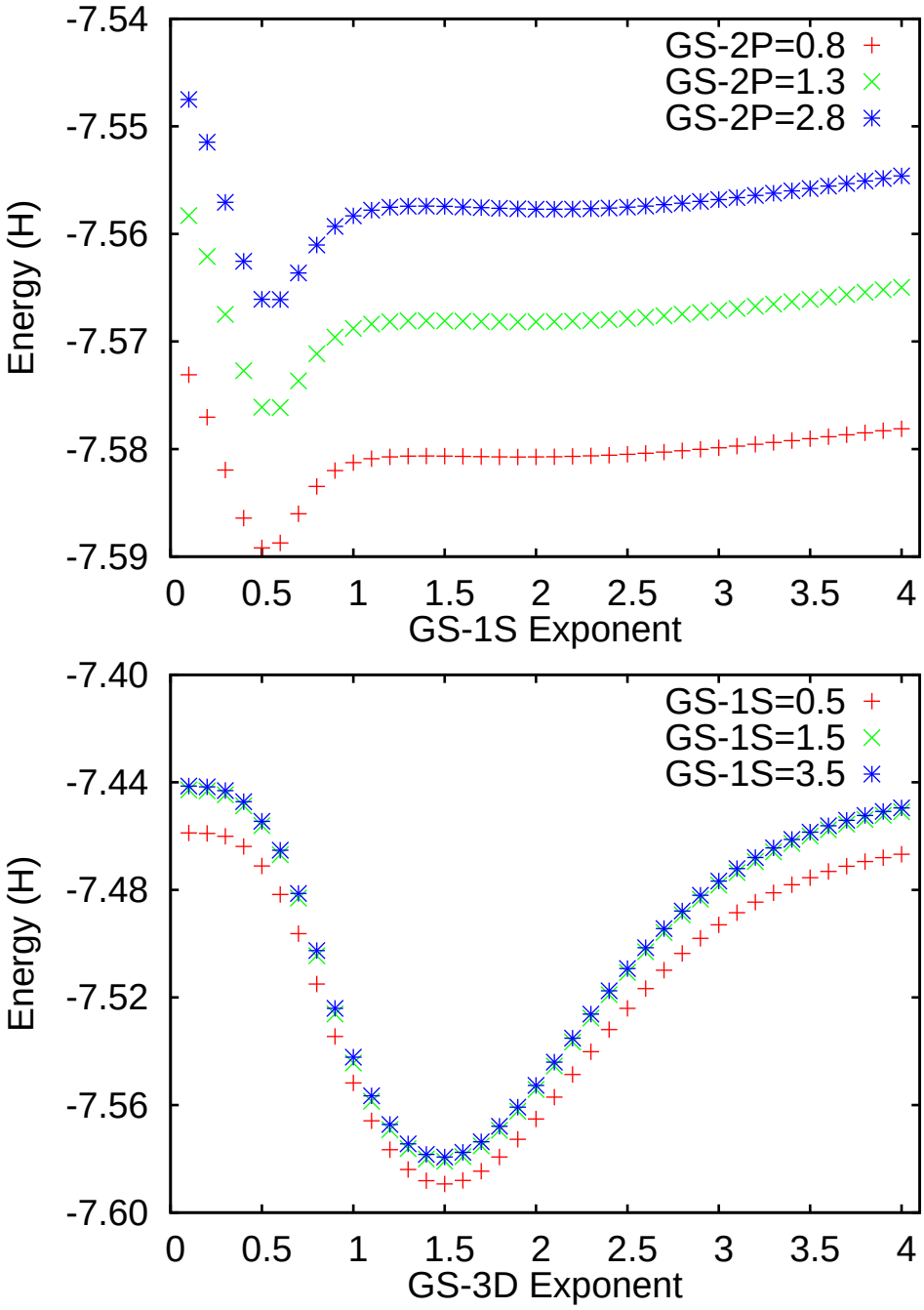

Figure 3.1: Change in Si_2 CCSD energy for $2z$ ANO-GS/GSn basis shows weak coupling between GS functions of different angular momenta. TOP: Energy versus GS-1S exponent for three values of the GS-2P exponent with the GS-3D exponent fixed at its optimal value. Bottom: Energy versus GS-3D exponent for three values of the GS-1S exponent with the GS-2P exponent fixed at its optimal value.

is defined by a grid of primitive exponents ranging from 0.1 to 6.0 with 0.1 spacing. Thorough investigation has revealed that exponents larger than 6.0 are not optimal for the systems considered. Low lying minima of this energy landscape are then handled with increasingly finer grids until energy changes are less than 0.01 mH. During this investigation of local minima, all angular momenta are handled simultaneously to account for any coupling effects. Results of this optimization are shown in Figure 3.2. Optimal exponents for ANO-GS and ANO-GSn bases exhibit a linear trend across each row of the periodic table. For nearly degenerate minima, the exponent following the trend in the figure is chosen as optimal, resulting in energy increase no greater than several 0.1 mH. The optimal GS exponents are given in the supplementary material [31] and Appendix C.

In some cases, the optimal exponents for primitives with the same n and l are very close. This can lead to large equal and opposite coefficients on these basis functions when constructing molecular orbitals. Numerical problems could result, providing further motivation for the ANO-GSn basis, in which each pair of n and l is unique. However, all of our tests with the ANO-GS basis have had no numerical problems.

Finally, the optimal primitive exponents are found to depend weakly on the electronic structure method employed in the optimization, as demonstrated in Figure 3.3 for Si_2 with the $2z$ ANO-GS/GSn basis. The globally minimizing exponents are nearly equal in different methods. This exponent transferability to different levels of theory is extremely attractive for a basis set.

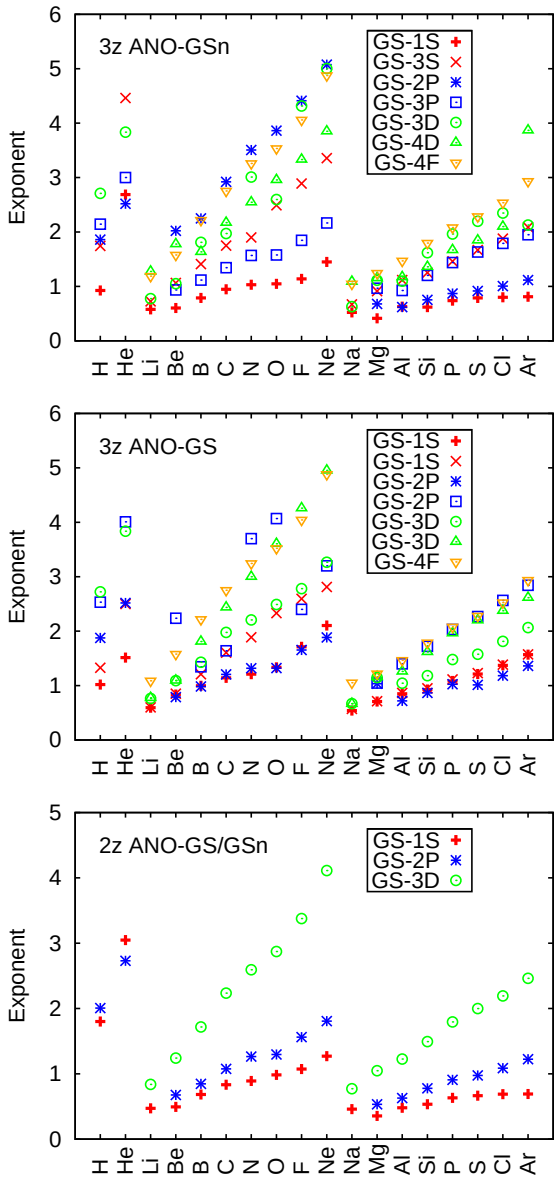

Figure 3.2: Optimal exponents for ANO-GS and ANO-GSn bases exhibit a linear trend across each row of the periodic table. The $2z$ ANO-GS and ANO-GSn bases are identical for all elements except lithium and sodium. The GS-1S and GS-3D exponents for these elements each differ by less than 0.01 between $2z$ ANO-GS and ANO-GSn bases, so $2z$ ANO-GS and ANO-GSn are shown together as $2z$ ANO-GS/GSn. Exponents for GS functions of P angular momentum are not included for lithium and sodium since these elements have an extra primitive of P angular momentum.

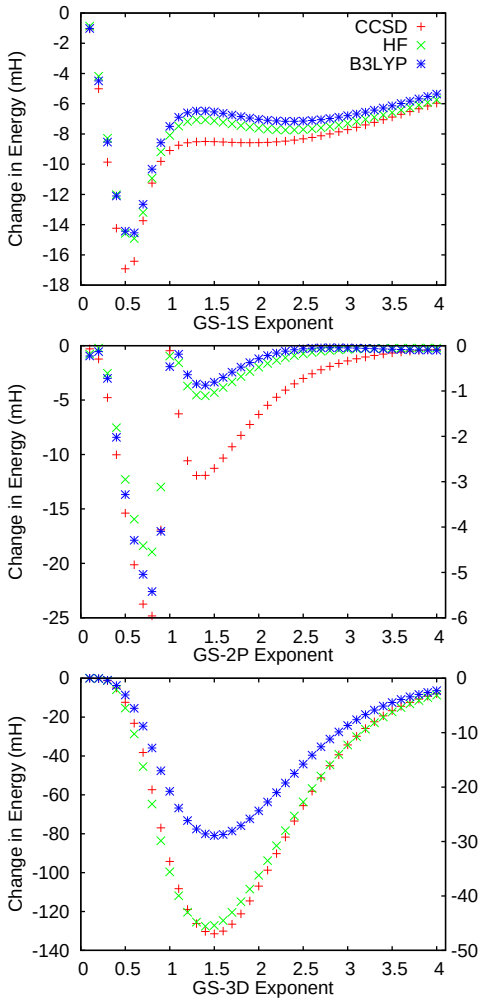

Figure 3.3: Change in Si_2 energy for $2z$ ANO-GS/GSn basis shows optimal exponents depend weakly on electronic structure method (CCSD, HF, and B3LYP). Top: GS-1S exponent is varied with GS-2P and GS-3D exponents fixed at their optimal values. Middle: GS-2P exponent is varied with GS-1S and GS-3D exponents fixed at their optimal values. The large increase in energy around an exponent of 1.0 occurs since the P primitive and P contraction become nearly linearly dependent. Bottom: GS-3D exponent is varied with GS-1S and GS-2P exponents fixed at their optimal values. For Middle and Bottom, HF and B3LYP energy scale is on the right y-axis. This difference in energy scale occurs since higher angular momentum functions are less important in these effectively single-determinant theories.

3.3 Results

Section 3.2 demonstrates that the ANO-GS and ANO-GSn bases exhibit desirable properties. However, it remains to be shown that these basis sets produce accurate results. Fortunately, the basis set accompanying the BFD pseudopotential serves as a metric for testing ANO-GS and ANO-GSn basis quality. The BFD basis for elements in Groups $1A$ and $2A$ of the periodic table has recently been updated [41], but the number of functions in the new basis is inconsistent with the correlation consistent polarized basis prescription [10]. Since comparison would be difficult, their published functions are considered in this work.

Figure 3.4 shows the CCSD total energy gain per electron of the ANO-GS and ANO-GSn bases over the BFD bases [18] for atoms and homonuclear dimers of hydrogen through argon. Energy gains per electron tend to increase across each row of the periodic table. Both ANO-GS and ANO-GSn bases yield energy gains for most molecules and atoms. The energy gains per electron are generally larger for molecules than for atoms, and larger for the ANO-GSn basis than for the ANO-GS basis. The energy gains for the $2z$ bases are generally larger than for the $3z$ bases, as expected, since the energy left to recover becomes smaller as the basis size increases.

The ANO-GS and ANO-GSn bases also produce more accurate CCSD atomization energies than the BFD basis for the homonuclear dimers of hydrogen through argon. Figure 3.5 shows the fraction of experimental atomization energy recovered in CCSD for the homonuclear dimers which are not weakly bound. The $2z$ ANO-GS/ANO-GSn basis recovers more atomization energy

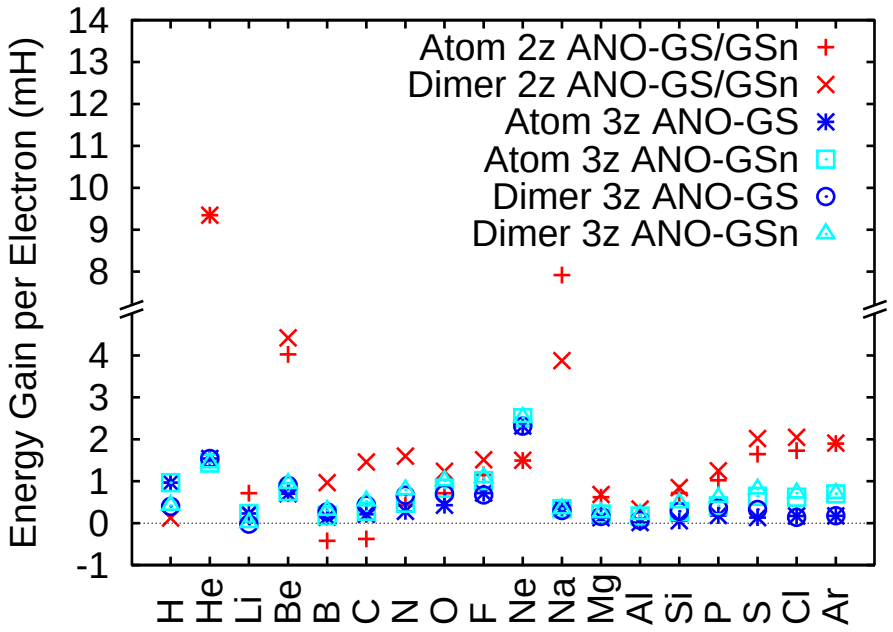

Figure 3.4: CCSD total energy gains per electron of ANO-GS and ANO-GSn relative to the corresponding BFD basis [18] for atoms and homonuclear dimers of hydrogen through argon. Energy gains per electron tend to increase across each row of the periodic table. The $2z$ ANO-GS and ANO-GSn bases are identical for all elements except lithium and sodium. Differences between $2z$ ANO-GS and ANO-GSn results for these elements is ~ 0.01mH, so they are shown together as $2z$ ANO-GS/GSn.

than the $2z$ BFD basis for all dimers except those of Group $1A$ elements. Similarly, the $3z$ ANO-GSn basis recovers more atomization energy than the $3z$ BFD basis for the same systems, but the differences are small. The $3z$ ANO-GSn is on average slightly better than the $3z$ ANO-GS basis, the largest gains being for F_2 and Cl_2.

For Group $1A$ elements, the BFD bases recover more atomization energy in CCSD than do their ANO-GS or ANO-GSn counterparts. This occurs due to inaccurate BFD energies for the atoms, as can be seen in Figure 3.4. However,

as described above, we used the published BFD bases for these elements rather than the updated BFD bases [41] to maintain consistency.

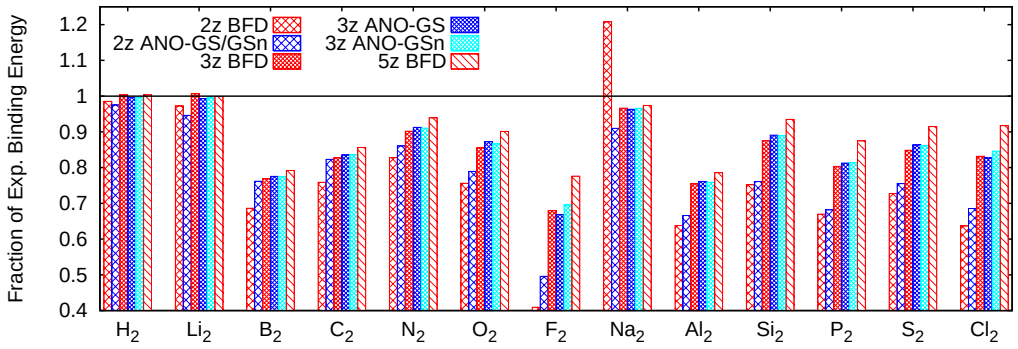

Figure 3.5: Fraction of experimental atomization energy recovered in CCSD with BFD, ANO-GS, and ANO-GSn bases for the homonuclear dimers of hydrogen through argon which are not weakly bound. The $2z$ ANO-GS and ANO-GSn bases are identical for all elements except lithium and sodium. Differences between $2z$ ANO-GS and ANO-GSn atomization energies for these elements is ~ 0.01mH, so they are shown together as $2z$ ANO-GS/GSn. Calculated values are corrected for zero point energy [42, 35] to compare with experiment [35, 32, 34, 43].

Finally, improvements of the ANO-GS and ANO-GSn bases extend to other systems and methods. Figure 3.6 shows the fraction of experimental atomization energy recovered for five systems in the G2 set [44] with the BFD, ANO-GS, and ANO-GSn bases in three quantum chemistry methods. For CCSD, the ANO-GS and ANO-GSn bases outperform the BFD basis for all systems. For sulfur dioxide the improvement due to the ANO-GS and ANO-GSn bases is dramatic: the $2z$ ANO-GS/GSn result is nearly halfway between the $2z$ and $3z$ BFD results, and the $3z$ ANO-GS/GSn result is nearly halfway between the $3z$ and $5z$ BFD results. ANO-GS and ANO-GSn benefits are more prominent in HF and B3LYP: for most systems, the $2z$ ANO-GS/GSn result is closer to the $3z$

BFD result than the $2z$ BFD result, and the $3z$ ANO-GS/GSn result is closer to the $5z$ BFD result than the $3z$ BFD result. Differences between results with the ANO-GS and ANO-GSn bases are small.

Figure 3.7 shows the fraction of experimental atomization energy recovered using diffusion Monte Carlo (DMC) with the BFD, ANO-GS, and ANO-GSn bases. For each system, the DMC calculations are performed with both a single-configuration state function (single-CSF) reference (DMC-1CSF) and full-valence complete active space reference (DMC-FVCAS). However, for each of the constituent atoms in these molecules, the FVCAS and single-CSF references are equivalent. All DMC calculations are performed with a 0.01 H^{-1} time step and trial wavefunction obtained by optimizing Jastrow, orbital, and configuration state function (CSF) parameters (where applicable) via the linear method [19, 20, 21] in variational Monte Carlo. The DMC-1CSF and DMC-FVCAS calculations exhibit similar trends to the HF and B3LYP calculation for most systems: the $2z$ ANO-GS/GSn result is closer to the $3z$ BFD result than the $2z$ BFD result, and the $3z$ ANO-GS/GSn result is closer to the $5z$ BFD result than the $3z$ BFD result. Again, differences between results with the ANO-GS and ANO-GSn bases are small.

There are several important points that can be made by comparing the DMC calculations of Figure 3.7 to the CCSD calculations of Figure 3.6. First, the DMC results for the atomization energies have a weaker dependence on basis size than the CCSD results. Second, for a given basis set, the most basic DMC calculations, DMC-1CSF, yield superior results compared to CCSD. In addition to yielding superior results, DMC-1CSF calculations have better computational cost scaling than CCSD calculations. Under certain assumptions, the cost of

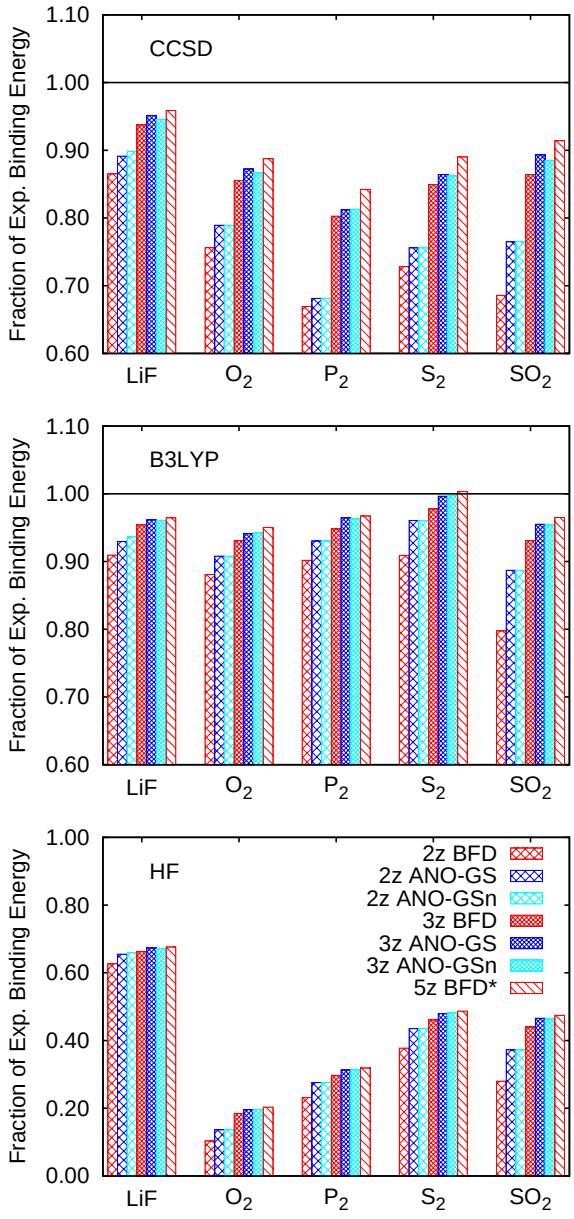

Figure 3.6: Fraction of experimental atomization energy recovered in HF, B3LYP, and CCSD for LiF, O_2, P_2, S_2, and SO_2 with BFD, ANO-GS, and ANO-GSn bases. The $2z$ ANO-GS and ANO-GSn bases yield different results only for LiF. The $5z$ BFD* calculations do not include the G or H functions from the $5z$ BFD basis. Calculated atomization energies are corrected for zero point energy [42, 35] to compare with experiment [35, 45, 32, 34].

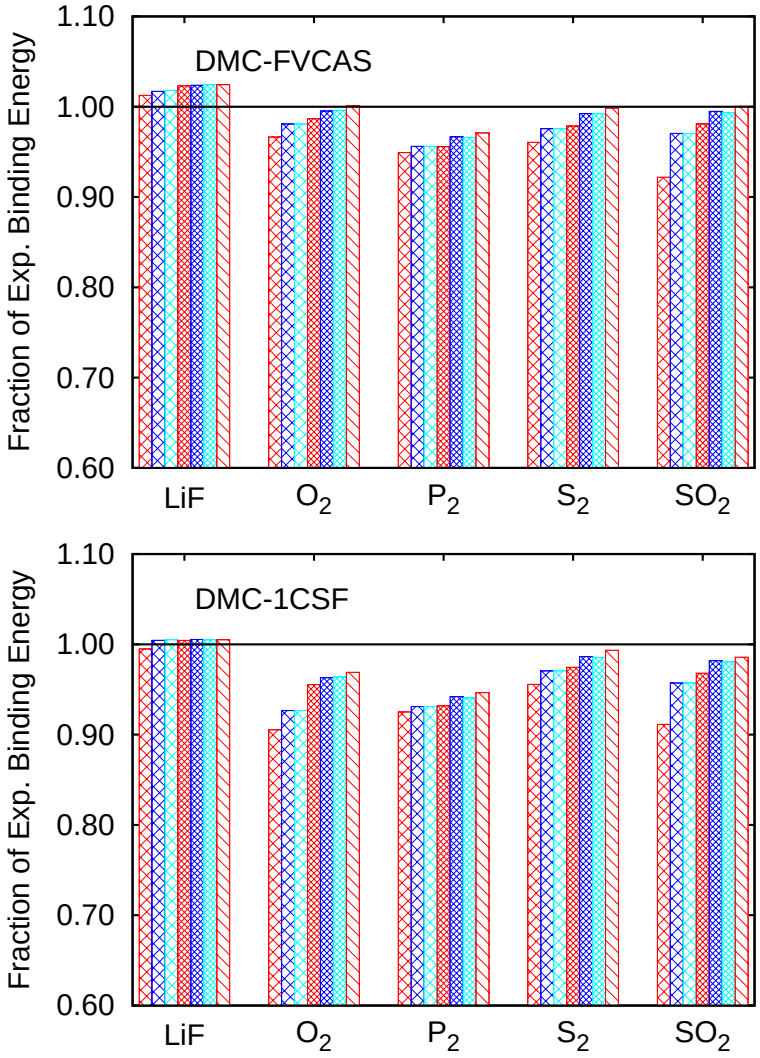

Figure 3.7: Fraction of experimental atomization energy recovered in diffusion Monte Carlo (DMC) for LiF, O_2, P_2, S_2, and SO_2 with the BFD, ANO-GS, and ANO-GSn bases. DMC calculations are performed with both a single-CSF reference (DMC-1CSF) and full-valence complete active space reference (DMC-FVCAS). The $2z$ ANO-GS and ANO-GSn bases yield different results only for LiF. The $5z$ BFD* calculations do not include the G or H functions from the $5z$ BFD basis. Calculated atomization energies are corrected for zero point energy [42, 35] to compare with experiment [35, 45, 32, 34]. The legend for this plot is identical to that of Figure 3.6.

DMC-1CSF calculations scales as $\mathcal{O}(N^3)$ [46], while the cost of CCSD calculations scales as $\mathcal{O}(N^6)$ [3], where N is the number of electrons. However, it is important to note that the prefactor of the scaling is significantly smaller for the CCSD calculations.

Finally, our results are not the first to show that DMC calculations can produce accurate atomization energies. In particular, DMC-1CSF calculations of the entire G2 set have been performed for both pseudopotential and all-electron systems [47, 48] and produced excellent results. Additionally, there is good agreement between the pseudopotential and all-electron results with a mean absolute deviation of about 2.0 kcal/mol over the entire G2 set [48]. Although these previous results are very good, there is room for improvement, particularly for the open shell systems. A systematic study with DMC-FVCAS calculations is currently underway in our group, which should produce results to (near) chemical accuracy for all systems in the G2 set.

3.4 Conclusion

A simple yet general method for constructing basis sets for molecular electronic structure theory calculations has been presented. These basis sets consist of a combination of atomic natural orbitals from an MCSCF calculation with primitive functions optimized for the corresponding homonuclear dimer. The functional form of the primitive functions is chosen to have the correct asymptotics for the nuclear potential of the system.

It was shown that optimal exponents of primitives with different angular momenta are weakly coupled. This enables efficient determination of opti-

mal exponents. Additionally, it was demonstrated that the particular electronic structure method employed in optimization has little effect on the optimal values of the primitive exponents.

Two sets of $2z$ and $3z$ bases, ANO-GS and ANO-GSn, appropriate for the Burkatzki, Filippi, and Dolg non-divergent pseudopotentials were constructed for elements hydrogen through argon. Since these pseudopotentials do not diverge at nuclei and have a Coulomb tail, GS functions are the appropriate primitives.

It was demonstrated that both ANO-GS and ANO-GSn basis sets offer significant gains over the Burkatzki, Filippi and Dolg basis sets for CCSD, HF, B3LYP [30], and QMC calculations. Improvements were observed in both total energies and atomization energies. The latter indicates that basis sets providing a balanced description of atoms and molecules were produced by using both the atom and the dimer in the optimization. On average, the ANO-GSn basis is slightly better than the ANO-GS basis, but either is a sound choice.

In the future, these basis sets will be extended to include the transition metals, and, bases will be constructed for all-electron calculations, for which Slater functions are the appropriate primitives.

CHAPTER 4

APPROACHING CHEMICAL ACCURACY WITH QUANTUM MONTE CARLO

The text of this chapter is a reproduction of a paper written on the same subject in 2012 [49]. The reference to the paper is J. Chem. Phys. 136, 124116 (2012). The abbreviations used in this chapter are given in Table 4.1.

Table 4.1: Abbreviations for Chapter 4. Descriptions are provided for non-standard abbreviations.

Abbreviation	Description
QMC	Quantum Monte Carlo
SJ	Slater-Jastrow: Product of Slater determinant and function which introduces correlation
MAD	Mean absolute deviation
HF	Hartree-Fock
DMC	Diffusion Monte Carlo
nz	$n-$zeta basis
VMC	Variational Monte Carlo
CAS	Complete Active Space
BFD	Burkatzki, Filippi and Dolg: Gaussian basis sets and non-divergent pseudopotentials constructed for QMC
ANO-GS	Basis with contractions that are atomic natural orbitals and primitives that are Gauss-Slaters
CSF	Configuration state function
CI	Configuration Interaction

4.1 Introduction

Quantum Monte Carlo (QMC) [4] is considered by some to be a "very accurate" method. However, previous QMC studies of the atomization energies of the molecules in the G2 set [44] have not obtained chemical accuracy [47, 48], defined as 1 kcal/mol. These studies, which are limited to a single determinant Slater-Jastrow (SJ) trial wavefunction and a fixed set of orbitals obtained via a quantum chemistry calculation, produce a mean absolute deviation (MAD) from experimental atomization energies of about 3 kcal/mol.

This work aims to improve upon both of those shortcomings. As a starting point, a single determinant SJ trial wavefunction composed of Hartree-Fock (HF) orbitals is used to compute the diffusion Monte Carlo (DMC) [16] atomization energies for the G2 set. These calculations, which are performed for double-zeta ($2z$), triple-zeta ($3z$), and quintuple-zeta ($5z$) bases, demonstrate the convergence of the DMC atomization energies with respect to basis size. The MAD from experiment for the $5z$ basis is 3.0 kcal/mol, in agreement with previous QMC studies [47, 48].

Next, the restriction to a fixed set of molecular orbitals is relaxed. The orbitals for each system and basis are optimized in variational Monte Carlo (VMC) via the linear method [19, 20, 21]. Employing the single determinant SJ trial wavefunction with optimized orbitals, DMC yields a MAD from experiment of 2.1 kcal/mol for the $5z$ basis.

Finally, the restriction of a single determinant SJ trial wavefunction is relaxed. With a complete active space (CAS) SJ trial wavefunction formed from just an s and p valence orbital active space, DMC produces atomization energies

of near chemical accuracy. The MAD from experimental atomization energies is 1.2 kcal/mol. This lends some backing to the claim that QMC is "very accurate". It is found that the MAD can be further reduced by including valence d orbitals in the active space for the heavier systems that are underbound in DMC.

This paper is organized as follows. In Section 4.2, the computational setup is described. In Section 4.3, results of the computations are described. Concluding remarks are in Section 4.4.

4.2 Computational Setup

All QMC calculations performed for this work use the Burkatzki-Filippi-Dolg (BFD) pseudopotentials [18, 41] in the QMC package CHAMP [17]. The $2z$ and $3z$ basis sets are the recently developed atomic natural orbital Gauss-Slater (ANO-GS) bases [7, 25]. For the $5z$ basis, the Gaussian BFD basis set [18, 41] is used, omitting the g and h functions. In the course of this study, it was determined that the hydrogen pseudopotential produced unreliable atomization energies. A significantly improved pseudopotential for hydrogen was developed by Filippi and Dolg, and is used in this work. Also, $2z$ and $3z$ ANO-GS basis sets, and a $5z$ Gaussian basis set appropriate for this pseudopotential have been constructed for this work. The improved pseudopotential is available upon request and the corresponding basis sets are available in the supplementary material [50] and Appendix D.

A combination of experimental and theoretical molecular geometries are used in this study [51, 52, 35, 53]. The zero point energies and experimental atomization energies are from Feller et al. [45, 51]. The geometries, zero point

energies, and experimental atomization energies for each molecule are available in the supplementary material [50] and Appendix D.

For single determinant SJ trial wavefunctions, the initial orbitals are generated in GAMESS [24] via spin-restricted Hartree-Fock calculations. The Jastrow parameters, and when applicable, the orbital parameters, are then optimized in VMC via the linear method [19, 20, 21].

For CAS SJ trial wavefunctions, the initial orbitals and initial configuration state function (CSF) coefficients are generated in GAMESS via multi-configurational self-consistent field theory (MCSCF) calculations. The Jastrow, orbital, and CSF parameters are then optimized in VMC via the linear method. The active space consists of the $1s$ orbital for hydrogen, the $2s$ and $2p$ orbitals for the first row atoms, and the $3s$ and $3p$ orbitals for the second row atoms, and the corresponding orbitals for the molecules.

Additionally, not all of the CSFs generated by the MCSCF calculations are included in the QMC calculations. Instead, a dual criterion for selecting CSFs is employed. If the magnitude of a CSF coefficient is at least 0.005 or a CSF is a double excitation from the HF CSF, then it is included in the trial wavefunction. This dual criterion is employed in contrast to the usual single criterion based only on the magnitude of CSF coefficients because the optimal CSF coefficients in QMC can differ greatly from the coefficients generated via MCSCF. Although the magnitude of most CSF coefficients decrease upon optimization in VMC due to the Jastrow factor's effectiveness in describing electronic correlations, there are systems for which the magnitude of the coefficients for a few double excitations increase considerably. This dual selection criterion results in a relatively modest number of CSFs. The largest number employed is for C_2H_6 and Si_2H_6.

These trial wavefunctions consist of 650 CSFs comprising 1700 unique determinants, whereas the MCSCF calculation generates 1.4 million CSFs.

Finally, all DMC calculations are performed with a 0.01 Hartree^{-1} time step. The walker populations are large enough for a negligible population control bias and furthermore the small population control bias is eliminated using the method described in Refs. [54, 16]. For all systems except LiH, BeH, CH_2 ($^3B_1$), LiF, C_2H_2, CN, HCN, HCO, NaCl the locality approximation [55] is employed for the nonlocal pseudopotential. The aforementioned systems suffer from instabilities with the locality approximation, so those computations are performed with the size-consistent version of the T-moves approximation [56]. Note that for these systems the atomic energies are also calculated with T-moves so that atomization energies are always calculated in a consistent manner. All DMC calculations are performed with a sufficient number of Monte Carlo steps such that the statistical error bar on the atomization energy of each system is about 0.1 kcal/mol.

4.3 Results

The raw data for all calculations presented here are available in the supplementary material [50] and Appendix D.

The deviation of the DMC atomization energies from experiment for a single determinant SJ trial wavefunction composed of HF orbitals is shown in Figure 4.1. The results for $2z$, $3z$, and $5z$ basis sets demonstrate the convergence of the atomization energies with respect to basis size. The MAD from experiment for the $2z$, $3z$, and $5z$ bases are 4.5 kcal/mol, 3.2 kcal/mol, and 3.0 kcal/mol,

respectively. The $5z$ result agrees with previous QMC studies [47, 48] which had a MAD from experiment of about 3 kcal/mol. Note that Nemec et. al performed all-electron DMC calculations with HF orbitals [48] whereas Grossman employed the Stevens-Basch-Krauss pseudopotentials [57] with MCSCF natural orbitals [47].

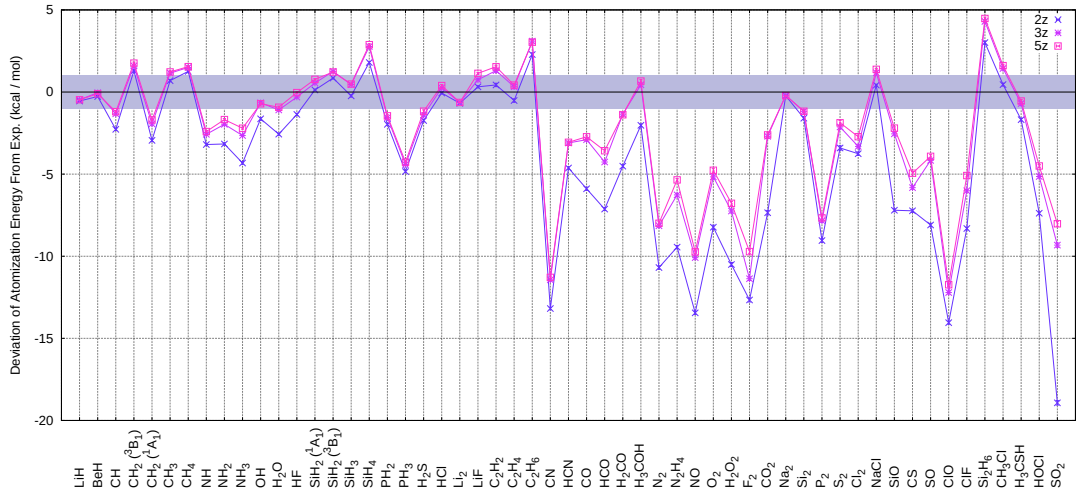

Figure 4.1: Deviation of the DMC atomization energies from experiment for a single determinant SJ trial wavefunction composed of HF orbitals. The MAD from experiment for the $2z$, $3z$, and $5z$ bases are 4.5 kcal/mol, 3.2 kcal/mol, and 3.0 kcal/mol, respectively.

Although orbitals from a quantum chemistry calculation are a reasonable starting point for a QMC calculation, they are certainly not optimal due to the presence of a Jastrow factor in the QMC wavefunction. Consequently, more accurate results are obtained by optimizing the orbitals in VMC. The deviation of the DMC atomization energies from experiment for a single determinant SJ trial wavefunction composed of VMC optimized orbitals is shown in Figure 4.2. Again, the results for $2z$, $3z$, and $5z$ basis sets demonstrate the convergence of the atomization energies with respect to basis size. The MAD from experiment

for the $2z$, $3z$, and $5z$ bases are 3.1 kcal/mol, 2.3 kcal/mol, and 2.1 kcal/mol, respectively.

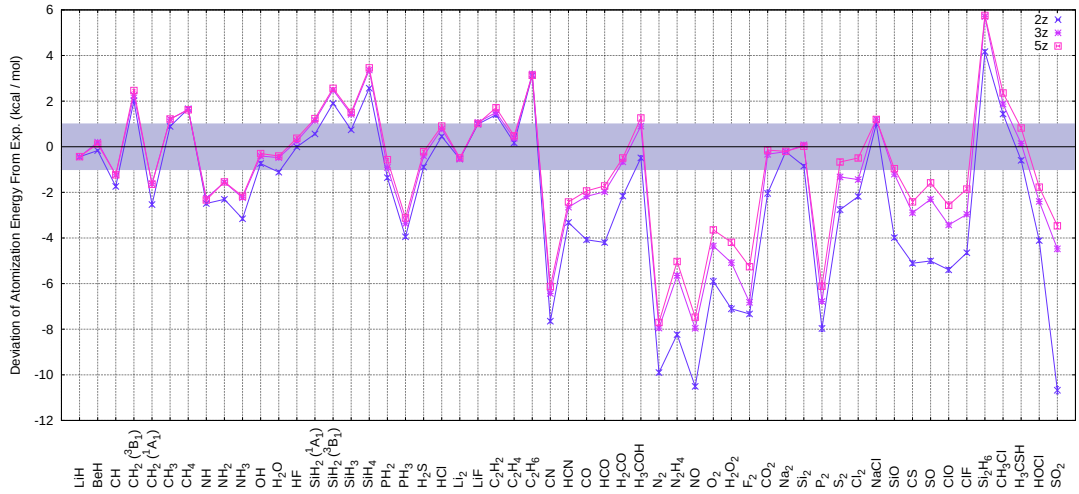

Figure 4.2: Deviation of the DMC atomization energies from experiment for a single determinant SJ trial wavefunction composed of VMC optimized orbitals. The MAD from experiment for the $2z$, $3z$, and $5z$ bases are 3.1 kcal/mol, 2.3 kcal/mol, and 2.1 kcal/mol, respectively.

As seen in Figure 4.3, the orbital optimized results are noticeably better than previous QMC studies [47, 48] which produce a MAD from experiment of about 3.0 kcal/mol. The gains in MAD from orbital optimization are 1.4 kcal/mol, 0.9 kcal/mol, and 0.9 kcal/mol for the three bases, respectively. Although, the largest gain is for the $2z$ basis, it is evident that the benefits of orbital optimization remain for even the largest basis set. It is worth pointing out that using optimized orbitals and a $2z$ basis produces results of similar quality to HF orbitals with a $5z$ basis.

Although orbital optimization provides significant improvements to the atomization energy, the results are still a long way off from chemical accuracy.

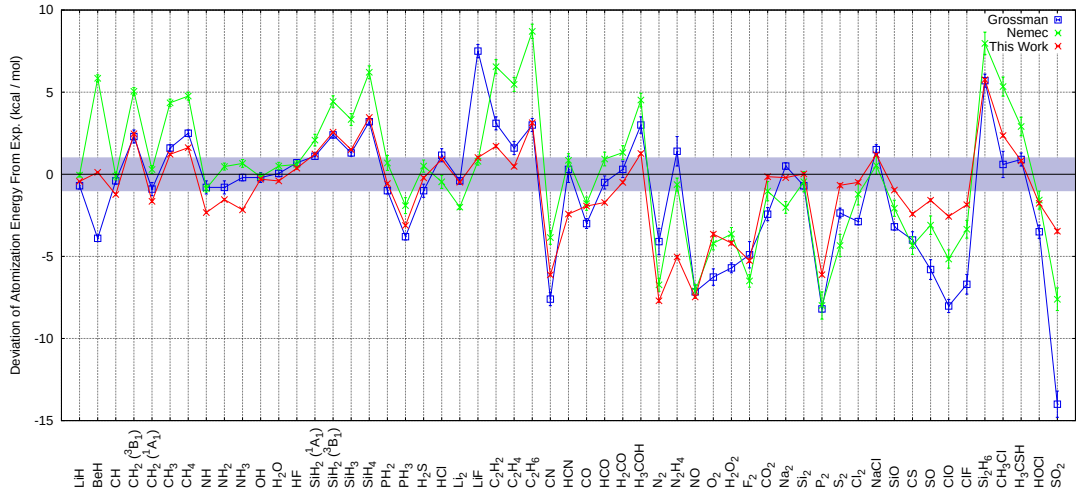

Figure 4.3: Comparison of the deviation of the DMC atomization energies from experiment for a single determinant SJ trial wavefunction. The results from this work are for a $5z$ basis and VMC optimized orbitals. The MAD from experiment for this work is 2.1 kcal/mol. The results of Nemec et al. and Grossman [47, 48] were obtained with HF orbitals and MCSCF natural orbitals, respectively. The MAD from experiment for Nemec et al. and Grossman are 3.1 and 2.9 kcal/mol, respectively.

To approach chemical accuracy, it is necessary to move beyond a single determinant SJ trial wavefunction because orbital optimization alone does not provide sufficient flexibility in the nodal surface of the trial wavefunction. Since the MAD of atomization energies from experiment for the $3z$ basis is only 0.2 kcal/mol higher than that of the $5z$ basis, and the cost of performing orbital optimization scales quadratically with the number of basis functions, the $3z$ basis used here represents a compromise between accuracy and computational efficiency. The deviation of the DMC atomization energies from experiment for the s and p valence CAS SJ trial wavefunctions is shown in Figure 4.4. The $5z$ single determinant results are included to demonstrate the benefit of using a CAS SJ trial wavefunction. This modest basis and CSF expansion results in a MAD

from experiment of 1.2 kcal/mol, a significant step forward for QMC.

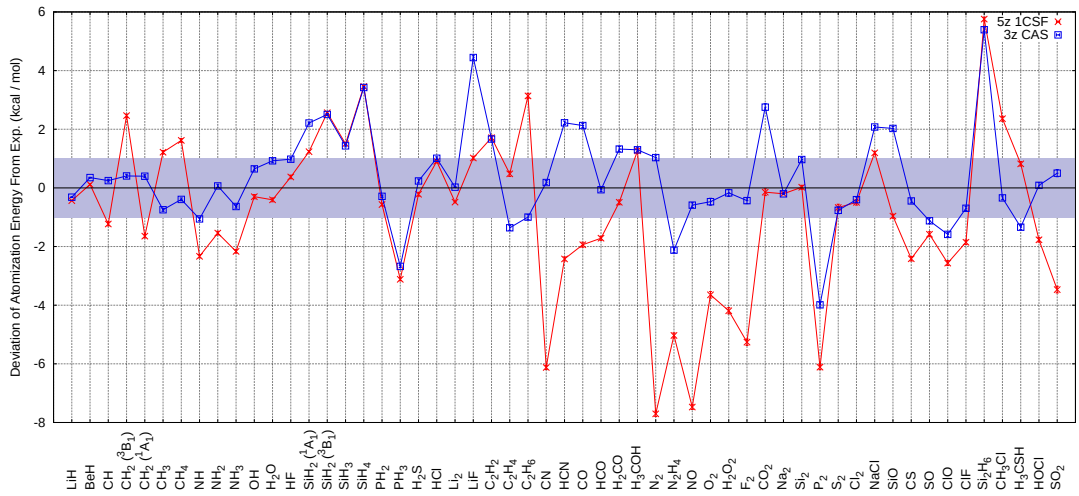

Figure 4.4: Deviation of the DMC atomization energies from experiment for a single determinant SJ trial wavefunction composed of VMC optimized orbitals and a CAS SJ trial wavefunction. The MAD from experiment for the single determinant SJ trial wavefunction is 2.1 kcal/mol. The MAD from experiment for the CAS SJ trial wavefunction is 1.2 kcal/mol.

As seen with the single determinant SJ results, both increasing the basis size and optimizing the orbitals have the effect of increasing the atomization energies for every system, since the energy gain is larger for the molecule than its constituent atoms. Since the small basis, single determinant SJ DMC results in most systems are underbound, this on average reduces the MAD of the atomization energies. On the other hand, going from single determinant to CAS trial wavefunctions increases the atomization energies for some systems and decreases it for others, but on average in the correct direction to reduce the MAD. For example, the atomization energies of CH and $CH_2(^1A_1)$ are increased and that of $CH_2(^3B_1)$ reduced, but all of these changes result in better agreement with experiment. However, using the CAS trial wavefunctions certainly does not always improve agreement with experiment, e.g. LiF and CO_2.

QMC can do yet better. Using a larger active space will certainly help, as the largest impediment for QMC is the fixed-node error. The choice of the modest s and p valence CAS allows for the possibility of scaling up to larger systems. However, for some systems an s and p valence CAS may not be sufficient to properly describe the nodal structure. To explore this, further study is performed on the phosphorous containing systems of the G2 set: PH_2, PH_3, P_2. Each of these systems is underbound for the s and p valence CAS. As shown in Figure 4.5, using s, p, and d valence CAS improves agreement between DMC atomization energies and experiment. The MAD from experiment for these three systems is 3.7, 2.3, and 1.6, for single determinant, s and p valence CAS, and s, p, and d valence CAS, respectively.

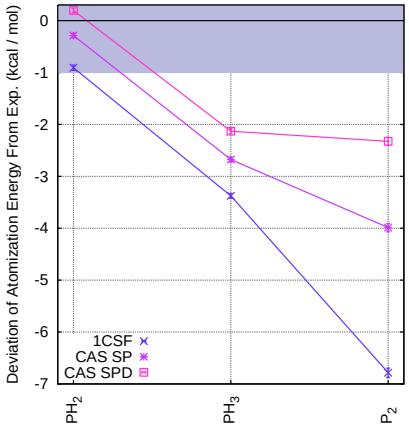

Figure 4.5: Deviation of the DMC atomization energies from experiment for a single determinant SJ trial wavefunction composed of VMC optimized orbitals, a CAS SJ trial wavefunction with an s and p active space, and a CAS SJ trial wavefunction with an s, p, and d active space. The MAD from experiment for the phosphorous containing systems of the G2 set with these trial wavefunctions is 3.7, 2.3, and 1.6, respectively.

Although using a larger active space for the phosphorus systems is beneficial, a large active space becomes impractical as system size increases. Even

though the number of CSFs included in a QMC calculation via the dual criterion described in Section 4.2 is much smaller than the total number of CSFs for a given active space, it is impractical to even perform the initial MCSCF calculation for large systems. Some options for alleviating this problem are obtaining the initial trial wavefunction from less expensive configuration interaction (CI) rather than MCSCF calculations, or, from restricted active space rather than complete active space calculations.

It is likely that some of the deviations of our results from experiment are due to using pseudopotentials. These deviations could be evaluated by performing a similar study with the all-electron couloumbic potential. However, there are some advantages to using pseudopotentials too. First, all-electron calculations for molecules containing second and higher row atoms are expensive. Second, it is possible that the fixed-node error for a given active space is larger for all-electron calculations. Finally, the use of pseudopotentials provides a simple way of including the scalar relativistic corrections.

Additionally, some of the deviations of our results from experiment are likely due to errors in the experimental atomization energies or zero point energies. In particular, as seen in Figure 4.4, systems containing both Si and H systematically overbind. Additionally, very accurate all-electron frozen-core coupled cluster calculations which produce sub-1 kcal/mol MAD from experiment for the G2 set [45] also systematically overbind these systems. In particular, Feller et al. overbinds $SiH_2(^1A_1)$, $SiH_2(^3B_1)$, SiH_3, SiH_4, Si_2H_6 by 1.3 kcal/mol, 1.1 kcal/mol, 0.2 kcal/mol, 1.6 kcal/mol, 3.5 kcal/mol, respectively.

4.4 Conclusion

A QMC study of the atomization energies for the G2 set of molecules was presented. Basis size dependence of DMC atomization energies was studied with a single determinant SJ trial wavefunction formed from HF orbitals. With the largest basis set, the mean absolute deviation from experimental atomization energies for the G2 set was found to be 3.0 kcal/mol, in agreement with previous QMC studies.

It was determined that optimizing the orbitals within VMC improved the agreement between DMC and experiment, reducing the mean absolute deviation to 2.1 kcal/mol. In fact, using optimized orbitals and a $2z$ basis produced results of similar quality to HF orbitals with a $5z$ basis.

Finally, DMC results for a CAS SJ trial wavefunction were near chemical accuracy with MAD from experimental atomization energies of 1.2 kcal/mol. Although a MAD of 1.2 kcal/mol is a significant step forward for QMC, comparison with all-electron frozen-core coupled cluster calculations, which produce sub-1 kcal/mol results for the G2 set [45, 51], demonstrates there is still room for improvement. Several directions for improving upon the current results are larger active spaces, backflow transformations [58], yet more accurate pseudopotentials, or all-electron calculations.

CHAPTER 5

SEMISTOCHASTIC PROJECTION

The text of this chapter is in preparation for submission. The abbreviations used in this chapter are given in Table 5.1.

Table 5.1: Abbreviations for Chapter 5. Descriptions are provided for non-standard abbreviations.

Abbreviation	Description
HF	Hartree-Fock
FCI	Full Configuration Interaction
QMC	Quantum Monte Carlo
SQMC	Semistochastic quantum Monte Carlo
FCIQMC	Full Configuration Interaction quantum Monte Carlo

5.1 Introduction

Given an $N \times N$ Hermitian matrix H, an eigenvector and its corresponding eigenvalue can be obtained in at least three ways: diagonalization, deterministic projection, or stochastic projection. Diagonalization provides all eigenvectors and eigenvalues, but requires $\mathcal{O}(N^2)$ storage. The memory cost limits diagonalization to $N \lesssim 10^5$ on a single processor. Deterministic projection, a particular case of iterative diagonalization such as the Lanczos method, provides the dominant eigenvector and eigenvalue of the projector, which is a function of H constructed such that the desired state of H is the dominant state of the projector. Such methods require $\mathcal{O}(N)$ storage and are limited to $N \lesssim 10^{10}$ on a single

processor. Finally, stochastic projection provides the dominant eigenvalue and a stochastic sampling of the dominant eigenvector of the projector, for yet larger N. It has been applied successfully for $N \sim 10^{108}$ [59][1].

Although stochastic projection can be applied to very large matrices, the eigenvalue obtained has a statistical uncertainty. For applications to matrices that have a sign problem[2], a bias is typically introduced to control the sign problem. In this work, an alternative method, semistochastic projection, is proposed for obtaining the dominant eigenvalue of a matrix. Like stochastic projection, this hybrid of deterministic and stochastic projection is applicable to large matrices. When compared to stochastic projection, the hybrid method significantly reduces the computational time required to obtain an eigenvalue with a specified level of uncertainty and often also reduces the bias for matrices with a sign problem.

Semistochastic projection has numerous potential applications: classical Monte Carlo[3], transfer matrix Monte Carlo [60], quantum Monte Carlo [1]. Semistochastic projection imposes two prerequisites. Matrices of interest have to be sparse, as is the case for stochastic projection. Secondly, one has to identify a relatively small subset of basis states that carry significant weight in the dominant eigenvector.

Both of these requirements are satisfied in many physical applications of interest. For example, the Hamiltonians for electronic structure problems are

[1]Note that the enormous space is somewhat misleading since most states are never visited. However, it demonstrates the effectiveness of stochastic projection in picking out important states when a hierarchy exists.

[2]A matrix is sign-problem free if the corresponding projector has rows and columns that have the same sign structure, aside from an overall sign. This ensures that all contributions to a given state are of the same sign.

[3]To find the dominant eigenvector of a Markov matrix.

sparse as they contain only one- and two-body operators. Additionally, the Hartree-Fock (HF) determinant, augmented by a small set of additional determinants, usually represent a significant fraction of the ground state wavefunction.

This work studies the application of semistochastic projection to finding the ground state energy of the quantum mechanical Hamiltonian in a discrete basis. In this context, deterministic projection is known as Full Configuration Interaction (FCI) to chemists and as exact diagonalization to physicists. Stochastic projection is the essence of various quantum Monte Carlo (QMC) methods. Accordingly, we refer to the method discussed here as semistochastic quantum Monte Carlo (SQMC).

There are several variants of QMC in a finite basis. Green's function Monte Carlo [61] has been applied to systems without a sign problem. Sign problems can be treated with fixed-node Green's function Monte Carlo [62]. However, the fixed-node approximation introduces an uncontrolled variational bias. Considerable effort has been expended in trying to reduce or eliminate the bias [63]. Recently, Full Configuration Interaction quantum Monte Carlo (FCIQMC) was introduced by Alavi and co-workers as an unbiased method for treating Hamiltonians with a sign problem [64]. However, practical implementations of FCIQMC use the *initiator* idea to control the sign problem [65][4].

[4]For initiator FCIQMC, there is a minimum weight threshold for spawning. Walkers that occupy a basis state which has total weight less than this threshold are not allowed to spawn onto unoccupied basis states. This increases the likelihood of annihilation and, hence, controls the sign problem. Initiator FCIQMC has a non-variational bias that vanishes in the limit of infinite walker population.

5.2 Theory: Semistochastic Projection

Given an $N \times N$ Hermitian matrix H in a finite, orthonormal basis $B = \{|\phi_1\rangle, \ldots, |\phi_N\rangle\}$, define the projector for the lowest eigenvalue E_0, and eigenvector $\psi^{(0)}$ with components $\psi_i^{(0)} \equiv \langle\phi_i|\psi^{(0)}\rangle$, as

$$P = 1 - \tau\left(H - E_T\right), \tag{5.1}$$

where $\tau < 2/(E_{\max} - E_0)$, E_T is a dynamically adjusted shift that approximates E_0, and $E_{\max}$ is the largest eigenvalue of H. Let

$$|\psi(0)\rangle = \sum_{i=1}^{N} w_i(0)|\phi_i\rangle, \tag{5.2}$$

be an arbitrary vector satisfying $\langle\psi(0)|\psi^{(0)}\rangle \neq 0$. Then, repeated application of P to $\psi(0)$ yields $\psi^{(0)}$,

$$P^M\psi(0) \propto \psi^{(0)}, \tag{5.3}$$

for sufficiently large M.

Projection methods thus far have used Eqn. (5.3) multiplying by P either deterministically or stochastically; semistochastic projection combines the two.

Define $\mathcal{D}$ as the set of deterministic indices and $\mathcal{S}$ as the set of stochastic indices, where $\mathcal{D} \cup \mathcal{S} = \{1, \ldots, N\}$, $\mathcal{D} \cap \mathcal{S} = \emptyset$, and $|D| \ll N$. Accordingly, P is the sum of a deterministic projector $P^{\mathcal{D}}$, and a stochastic projector $P^{\mathcal{S}}$,

$$P = P^{\mathcal{D}} + P^{\mathcal{S}}, \tag{5.4}$$

where

$$P_{ij}^{\mathcal{D}} = \begin{cases} P_{ij}, & i,j \in \mathcal{D}, \\ 0, & \text{otherwise}, \end{cases} \tag{5.5}$$

and $P^{\mathcal{S}} \equiv P - P^{\mathcal{D}}$.

Semistochastic projection involves a combination of deterministic application of $P^{\mathcal{D}}$ and stochastic application of $P^{\mathcal{S}}$. To understand the semistochastic projection algorithm, it is sufficient to consider how an arbitrary vector at projection time $n\tau$,

$$|\psi(n\tau)\rangle = \sum_{i=1}^{N} w_i(n\tau)|\phi_i\rangle, \tag{5.6}$$

evolves to time $(n+1)\tau$. The coefficients of the basis functions are represented as a population of *walkers*, where the number of walkers on an occupied $|\phi_i\rangle$ is

$$n_i = \text{min}(1, \lfloor |w_i| \rceil), \tag{5.7}$$

where $\lfloor \cdot \rceil$ denotes the nearest integer and each walker has weight w_i/n_i. For semistochastic projection, the basis coefficients are evolved to time $(n+1)\tau$ according to the following algorithm.

- To account for the off-diagonal elements in $P^{\mathcal{S}}$, for each walker on $|\phi_i\rangle$, a move to $|\phi_j\rangle \neq |\phi_i\rangle$ is proposed with probability T_{ji}, where T is a stochastic matrix called the proposal matrix. The magnitude of the contribution to the walker weight on $|\phi_j\rangle$ from a single walker on $|\phi_i\rangle$ is

$$\begin{cases} 0, & i,j \in \mathcal{D} \\ \frac{P_{ji}}{T_{ji}} \frac{w_i(n\tau)}{n_i(n\tau)} = -\tau \frac{H_{ji}}{T_{ji}} \frac{w_i(n\tau)}{n_i(n\tau)} & \text{otherwise} \end{cases}. \tag{5.8}$$

  If H has a sign problem, then it is necessary to use the initiator idea [65]. Here it is generalized, in that the initiator criterion only applies to basis states outside of the deterministic space.

- To account for the diagonal elements in $P^{\mathcal{S}}$, the contribution to the total

walker weight on $|\phi_j\rangle$, with $j \in \mathcal{S}$, is

$$P_{jj} w_j(n\tau) = [1 - \tau(H_{jj} - E_T)]\, w_j(n\tau). \tag{5.9}$$

- Deterministic evolution is performed with $P^{\mathcal{D}}$. The contribution to the weight on $|\phi_j\rangle$, with $j \in \mathcal{D}$, is

$$\sum_{i \in \mathcal{D}} P^{\mathcal{D}}_{ji} w_i(n\tau). \tag{5.10}$$

  $P^{\mathcal{D}}$ is stored and applied as a sparse matrix.

- Finally, for each $|\phi_j\rangle$, all walker weight generated on $|\phi_j\rangle$ is summed, taking into account the sign of the contribution. To avoid the large computational and memory cost of having small weights on a large number of basis states, basis states with weights less than some minimum cutoff, w_{min}, are combined via an unbiased prescription [16].

After sufficiently many applications of the projector, contributions from subdominant eigenvectors die out. At this point, the collection of averages begins. The most commonly employed estimator is the mixed estimator for the dominant eigenvalue,

$$E_{\text{mix}} = \frac{\langle \psi^{(0)} | \hat{H} | \psi_T \rangle}{\langle \psi^{(0)} | \psi_T \rangle}, \tag{5.11}$$

where the trial state $|\psi_T\rangle$ satisfies $\langle \psi^{(0)} | \psi_T \rangle \neq 0$.

The trial state $|\psi_T\rangle$ is a linear combination of basis states[5]

$$|\psi_T\rangle = \sum_{i \in \mathcal{T}} d_i |\phi_i\rangle, \tag{5.12}$$

where $\mathcal{T}$ is the set of indices corresponding to those basis functions which have non-zero coefficient in the trial state, and $|\mathcal{T}| \ll N$. It is not necessary that $\mathcal{T} \subset \mathcal{D}$.

[5] In FCIQMC [59, 64, 65], a single state, the HF determinant, has been used as the trial state.

At any particular time $n\tau$, the stochastic representation of the dominant eigenvector is

$$|\psi^{(0)}(n\tau)\rangle = \sum_{i \in \mathcal{W}(n\tau)} w_i(n\tau)|\phi_i\rangle, \tag{5.13}$$

where $\mathcal{W}(n\tau)$ is the set of indices corresponding to basis functions occupied by walkers at time $n\tau$. The full representation of the dominant eigenvector is obtained by summing over Monte Carlo generations

$$|\psi^{(0)}\rangle = \frac{1}{N_{\text{gen}}} \sum_{n=0}^{N_{\text{gen}}-1} \sum_{i \in \mathcal{W}(n\tau)} w_i(n\tau)|\phi_i\rangle, \tag{5.14}$$

where N_{gen} is the number of times the projector is applied after equilibration.

For the trial state in Eqn. (5.12),

$$E_{\text{mix}} = \frac{\sum_{n=0}^{N_{\text{gen}}-1} \sum_{i \in \mathcal{W}(n\tau)} w_i(n\tau) \sum_{j \in \mathcal{T}} H_{ij} d_j}{\sum_{n=0}^{N_{\text{gen}}-1} \sum_{i \in \mathcal{W}(n\tau) \cap \mathcal{T}} w_i(n\tau) d_i}. \tag{5.15}$$

Since E_{mix} is a zero-variance-zero-bias estimator when $|\psi_T\rangle$ is equal to the dominant eigenvector, improving the quality of $|\psi_T\rangle$ reduces fluctuations and bias in the mixed estimate of the dominant eigenvalue. This reduction can be achieved with almost no additional computational cost by storing nonzero $\sum_{j \in \mathcal{T}} H_{ij} d_j$ terms.

5.3 Applications

Semistochastic projection is now applied to finding the ground state energy of the quantum mechanical Hamiltonian in a discrete basis. The benefits of SQMC are demonstrated by comparing the efficiency of SQMC relative to a purely stochastic method for $SiH_2(^1A_1)$ and the 2-dimensional 8×8 Hubbard model. The stochastic method is SQMC with a deterministic space of size 1.

The efficiency is defined to be proportional to the inverse of the time required to obtain the ground state energy to a specified level of uncertainty.

For the following calculations, the trial wavefunction is chosen to be the HF determinant. The deterministic space is generated by picking a reference state, and then including all determinants connected to that reference state by a single application of the Hamiltonian. At a minimum, the reference includes the HF determinant, but generally includes a short determinant expansion.

Figure 5.1 shows the efficiency gains of SQMC relative to the stochastic method for $SiH_2(^1A_1)$ with a cc-pVDZ basis set [10, 66]. For the largest deterministic space considered, SQMC is over 100 times more efficient than the purely stochastic algorithm.

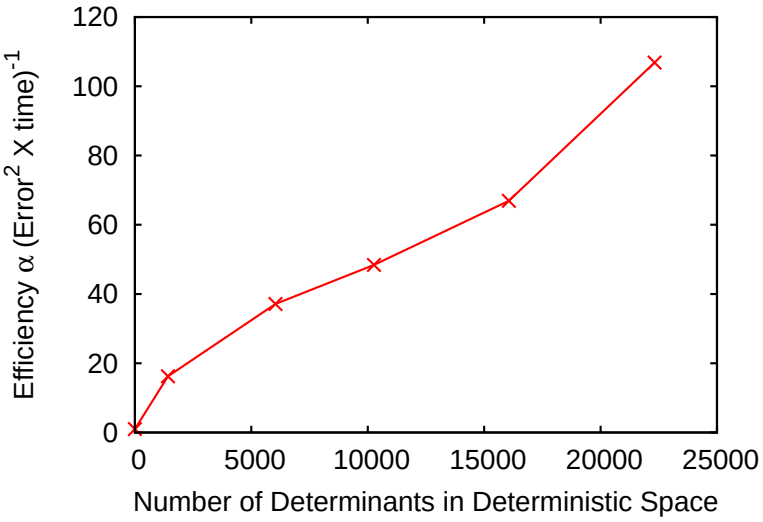

Figure 5.1: Efficiency of SQMC relative to the stochastic method for $SiH_2(^1A_1)$ with a cc-pVDZ basis set [10, 66]. The efficiencies are normalized by that of the stochastic method, which is shown as a deterministic space of size 1. These calculations are performed with the frozen-core approximation. The Hilbert space is roughly 10^6 determinants.

Figure 5.2 shows the efficiency gains of SQMC relative to the stochastic method for the two-dimensional 8×8 Hubbard model with $U/t = 4$ at several filling fractions. For 26 electrons, the efficiency gain is over 50. In fact, the efficiency gains increase with increasing filling fraction. This demonstrates the potential for SQMC to study strongly correlated systems.

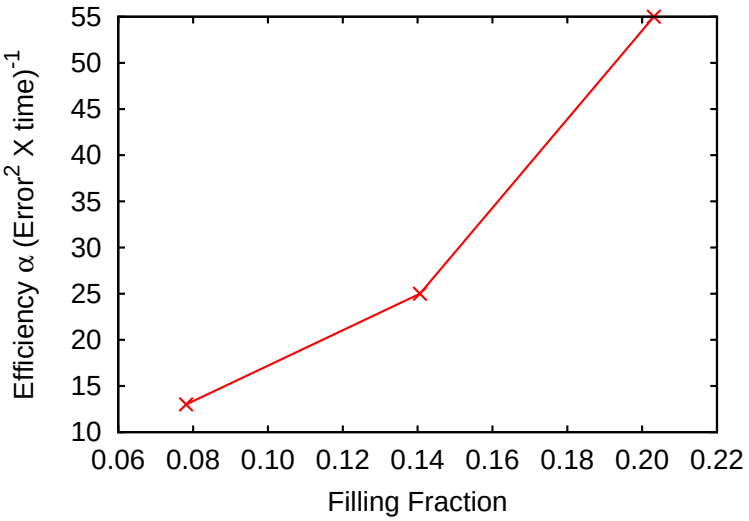

Figure 5.2: Efficiency of SQMC relative to the stochastic method at the same filling for the two-dimensional 8×8 Hubbard model with $U/t = 4$ at several filling fractions. The efficiencies at each filling are normalized by that of the stochastic method. The Hilbert spaces range from 10^{12} to 10^{24}, far larger than can be handled with exact diagonalization. The deterministic space reference for each of these calculations is the HF determinant.

Not only does SQMC have large efficiency gains relative to the stochastic method, but in some cases, SQMC has a significantly reduced initiator bias. The initiator bias for $SiH_2(^1A_1)$ is extremely small, so only Hubbard model results are included. Figure 5.3 shows both the SQMC and stochastic method energy as a function of the average number of occupied determinants for 8×8 Hubbard model with $U/t = 1$ and 50 electrons. For all but the smallest number of

occupied determinants, SQMC has essentially no bias.

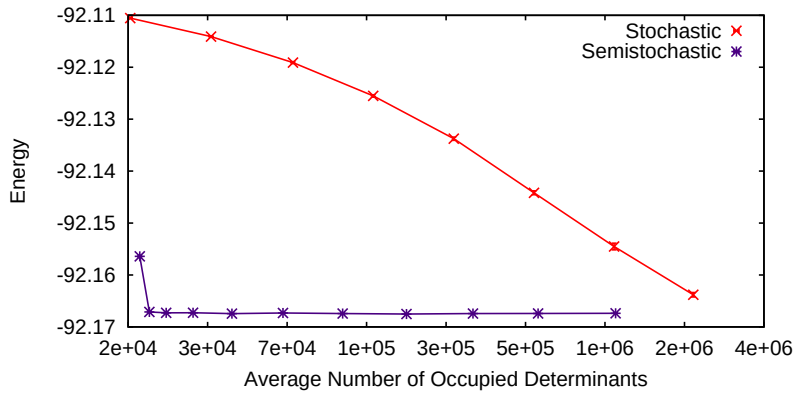

Figure 5.3: Energy of SQMC and the stochastic method vs. the average number of occupied determinants for the two-dimensional 8×8 Hubbard model with $U/t = 1$ and 50 electrons. The Hilbert space is about 10^{35}. The deterministic space reference for each SQMC calculation is the HF determinant, yielding a deterministic space of 16540 determinants.

However, the reduction in initiator bias is not always large. Figure 5.4 shows both the SQMC and stochastic method energy as a function of the average number of occupied determinants for 8×8 Hubbard model with $U/t = 4$ and 10 electrons. SQMC has a reduced initiator bias for small number of occupied determinants but not for large in this case.

5.4 Conclusion

Semistochastic projection, a hybrid between deterministic projection and stochastic projection, was introduced for finding the dominant eigenvalue and sampling the corresponding eigenvector of a matrix. It was demonstrated that, like stochastic projection, this method is applicable to matrices well beyond the

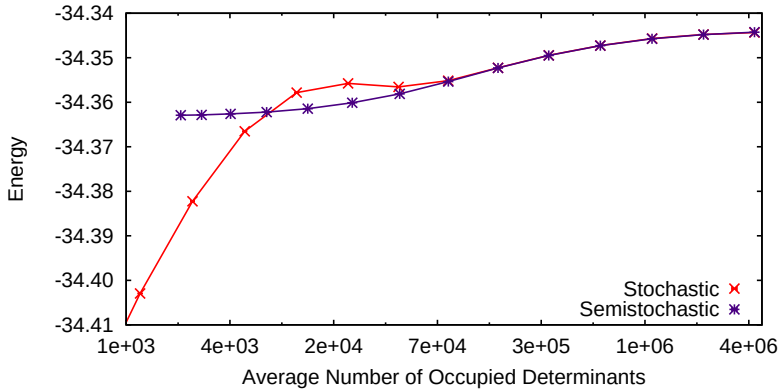

Figure 5.4: Energy of SQMC and the stochastic method vs. the average number of occupied determinants for the two-dimensional 8×8 Hubbard model with $U/t = 4$ and 10 electrons. The Hilbert space is about 10^{12}. The deterministic space reference for each SQMC calculation is the HF determinant, yielding a deterministic space of 1412 determinants.

size that can be handled by deterministic methods. In particular, Hilbert spaces ranging from 10^{12} to 10^{35} were tackled for the two-dimensional 8×8 Hubbard model. It was shown that semistochastic projection is far more efficient than stochastic projection and that in some cases it has the additional benefit of a much reduced initiator bias.

CHAPTER 6

CONCLUDING REMARKS

This thesis detailed a range of accomplishments that move the field of quantum Monte Carlo forward.

Chapter 2 focused on improving a specific aspect of the Slater-Jastrow wavefunction used in continuum quantum Monte Carlo calculations for molecular systems. In particular, we studied the single-particle basis functions employed and demonstrated that it is important for their functional form to be capable of producing the correct short- and long-range asymptotic behavior of the wavefunction. For calculations requiring non-diverging pseudopotentials, we introduced Gauss-Slater basis functions, which behave like Gaussians at short distances and Slaters at long distances. It was demonstrated that this form of basis function produces superior accuracy and efficiency when compared to the Gaussian basis sets typically employed in quantum Monte Carlo. By using Gauss-Slater functions, researchers in the field can employ a more compact basis, and hence, study larger systems.

Chapter 3 also focused on single-particle basis functions. Although we previously demonstrated that Gauss-Slater functions are capable of producing superior results, the parameters of these functions needed to be optimized for each system studied. Consequently, we introduced a simple, yet general, method for constructing basis sets appropriate for molecular electronic structure calculations. With this method, we created a standard Gauss-Slater basis set for each of the elements from hydrogen to argon. This contribution allows researchers in the field to employ the accurate Gauss-Slater functions in a black-box way just like standard Gaussian basis sets.

Chapter 4 explored the effect of different aspects of the trial wavefunction on the accuracy of quantum Monte Carlo calculations. By systematically testing the effect of the basis size, orbital quality, and determinant expansion quality, this work offers guidance to quantum Monte Carlo practitioners for achieving results to within chemical accuracy of experiment. Additionally, by following the guidance offered in this paper, these results can be achieved in a black-box way. This moves the field forward by encouraging non-experts to add quantum Monte Carlo to their computational physics/chemistry toolkit.

Chapter 5, introduced semistochastic projection for finding the dominant eigenvalue and eigenvector of matrices too large to be handled by deterministic projection. This hybrid between deterministic projection (exact diagonalization) and stochastic projection (quantum Monte Carlo) was demonstrated to be orders of magnitude more efficient than stochastic projection. Among other things, this development is a large step towards being able to study exotic physics in model systems like the Hubbard Model.

APPENDIX A

APPENDIX FOR CHAPTER 1

A.1 Metropolis-Hastings Algorithm

A.1.1 Foundations

Before we discuss the Metropolis-Hastings Algorithm it is necessary to lay the foundation by discussing the basics of Markov Chains. Markov Chains are a stochastic process where future states depend only on the present state. In other words, given some initial distribution, $\boldsymbol{\rho^0}$ and a stochastic matrix (also known as Markov matrix), M, the future distribution is completely determined by the application of M to $\boldsymbol{\rho^0}$.

A stochastic matrix is a matrix, M, satisfying the following

$$0 \leq M_{ij} \leq 1 \tag{A.1}$$

$$\sum_i M_{ij} = 1. \tag{A.2}$$

Note, a stochastic matrix, M, has at least one right eigenvector $\boldsymbol{\psi^0}$ with unit eigenvalue, $\lambda_0 = 1$. To see this, consider the vector $\boldsymbol{l} = (1, 1, \ldots, 1)$. It is clear the $\boldsymbol{l}^\mathsf{T}$ is a left eigenvector of M with eigenvalue 1,

$$\begin{aligned}
(\boldsymbol{l}^\mathsf{T} \cdot M)_j &= \sum_i l_i^\mathsf{T} M_{ij} \\
&= \sum_i M_{ij} \\
\text{[Eqn. (A.2)]} \qquad &= 1 \\
&= (1) l_j^\mathsf{T}.
\end{aligned} \tag{A.3}$$

Since left and right eigenvalues are equivalent, there must exist a right eigenvector, $\boldsymbol{\psi^0}$, with eigenvalue $\lambda_0 = 1$. It follows that any right eigenvector, $\boldsymbol{\psi^n}$, with eigenvalue $\lambda_n \neq 1$ must have components that sum to 0,

$$\begin{aligned}\sum_i \lambda_n \psi_i^n &= \sum_i \sum_j M_{ij} \psi_j^n \\ &= \sum_j \psi_j^n \left(\sum_i M_{ij} \right) \\ &= \sum_j \psi_j^n, \end{aligned} \tag{A.4}$$

so that for $\lambda_n \neq 1$, we have

$$\sum_j \psi_j^n = 0. \tag{A.5}$$

In other words, for $\lambda_n \neq 1$, $\boldsymbol{\psi^n}$ is not positive (has some negative components).

Now, we need several definitions and a very important theorem before we make any conclusions. A regular stochastic matrix, is a stochastic matrix, M, such that there exists a finite k where $M_{ij}^k > 0 \, \forall \, i, j$. A cyclic stochastic matrix, is a stochastic matrix, M, such that the state space can be divided into subspaces where there is a non-zero probability only for transitions from states in one subspace to states of the next subspace. An ergodic stochastic matrix, is a stochastic matrix, M, that is regular and not cyclic.

The Perron-Frobenius Theorem states that if a matrix, M, is ergodic, then there is a unique (up to multiplicative factor) positive eigenvector with a corresponding unique positive eigenvalue that has the greatest norm of all the eigenvalues of M (known as dominant eigenvalue). Hence, for an ergodic stochastic matrix, M, there must be a positive eigenvector. However, if $\lambda_n \neq 1$ then the corresponding eigenvector is non-positive. Thus, the unique eigenvalue and corresponding unique eigenvector must be $\lambda_0 = 1$ and $\boldsymbol{\psi^0}$, respectively.

The Perron-Frobenius theorem requires that all other eigenvalues of M have norm less than 1. Now, we are interested in the action of M on some initial distribution $\boldsymbol{\rho^0}$. As long as the right eigenvectors of M are complete, any initial distribution, $\boldsymbol{\rho^0}$, can be expanded in a linear combination of those eigenvectors. If $\boldsymbol{\rho^0}$ has some component along $\boldsymbol{\psi^0}$ then repeated application of M to $\boldsymbol{\rho^0}$ will project out $\boldsymbol{\psi^0}$,

$$\begin{aligned} M^k \cdot \boldsymbol{\rho^0} &= M^k \cdot \sum_n c_n \boldsymbol{\psi^n} \\ &= \sum_n c_n \lambda_n^k \boldsymbol{\psi^n} \\ (k \gg 1, \lambda_0 = 1 > |\lambda_n| \,\forall\, n \neq 0) \qquad &\approx c_0 \boldsymbol{\psi^0}. \end{aligned} \tag{A.6}$$

Notice that this implies that M^k has one eigenvalue of 1 and all other eigenvalues 0. It is clear that the magnitude of the sub-dominant eigenvalues, λ_n for $n > 0$, determine how large k must be for Eqn. (A.6) to be valid. In the ideal limit, $\lambda_n = 0$ for $n > 0$, $k = 1$ and the evolution from $\boldsymbol{\rho^0}$ to $\boldsymbol{\psi^0}$ takes only a single application of M. Notice the trivial result

$$M \cdot \boldsymbol{\psi^0} = \boldsymbol{\psi^0}. \tag{A.7}$$

This is known as the stationarity condition and implies that once we have the distribution $\boldsymbol{\psi^0}$ we will continue to have that distribution.

A.1.2 Choosing The Appropriate Markov Matrix

The goal of repeated application of an ergodic stochastic matrix, M, to a probability density $\boldsymbol{\rho^0}$ is to project out a target probability distribution, $\boldsymbol{\psi^t}$. How does one pick an appropriate M to produce $\boldsymbol{\psi^t}$? For this, the detailed balance

condition is usually imposed

$$M_{ij}\psi_j^t = M_{ji}\psi_i^t. \tag{A.8}$$

In this case we are guaranteeing that the stationary distribution is the target distribution, $\boldsymbol{\psi^t}$,

$$\begin{aligned} \left[M \cdot \boldsymbol{\psi^t}\right]_i &= \sum_j M_{ij}\psi_j^t \\ \text{[Eqn. (A.8)]} \qquad &= \sum_j M_{ji}\psi_i^t \\ \text{[Eqn. (A.2)]} \qquad &= \psi_i^t \end{aligned} \tag{A.9}$$

Now, we write the elements of M as a product of elements of a proposal matrix, T, and an acceptance matrix, A,

$$M_{ij} = \begin{cases} A_{ij}T_{ij} & i \neq j \\ 1 - \sum_{k\neq j} A_{kj}T_{kj} & i = j, \end{cases}$$

Note that we choose T as a stochastic matrix by construction. We also choose A to have non-negative elements. It follows immediately that M is a stochastic matrix

$$\begin{aligned} \sum_i M_{ij} &= 1 - \sum_{k\neq j} A_{kj}T_{kj} + \sum_{i\neq j} A_{ij}T_{ij} \\ &= 1 \end{aligned} \tag{A.10}$$

The acceptance matrix, A, is chosen so that detailed balance is satisfied. Note that detailed balance is automatically satisfied for diagonal elements of A so upon considering the off-diagonal elements we have

$$A_{ij}T_{ij}\psi_j^t = A_{ji}T_{ji}\psi_i^t. \tag{A.11}$$

which implies that

$$\frac{A_{ij}}{A_{ji}} = \frac{T_{ji}\psi_i^t}{T_{ij}\psi_j^t}. \tag{A.12}$$

There are an unlimited number of choices for the acceptance matrix that satisfy Eqn. (A.12). In fact, any function

$$A_{ij} = f\left(\frac{T_{ji}\psi_i^t}{T_{ij}\psi_j^t}\right) \tag{A.13}$$

for which $f(x)/f(1/x) = x$ will work. The choice corresponding to the Metropolis-Hastings method is

$$A_{ij} = \min\left(1, \frac{T_{ji}\psi_i^t}{T_{ij}\psi_j^t}\right). \tag{A.14}$$

The choice corresponding to the simple Metropolis method, which involves a symmetric proposal matrix, is

$$A_{ij} = \min\left(1, \frac{\psi_i^t}{\psi_j^t}\right). \tag{A.15}$$

Finally, the major advantage of using the Metropolis-Hastings method is that picking a clever proposal matrix, T, can result in large efficiency gains [67, 1].

A.1.3 Example Proposal Matrix

In this section we derive a proposal matrix, T, via the Fokker-Plank formalism. Note that this is equivalent to taking the importance sampled Schrodinger equation and setting the branching term to 0. This allows for one to share much of the same code for VMC and DMC.

Consider the Fokker-Plank PDE

$$\frac{\partial f}{\partial t} = D\nabla^2 f - \boldsymbol{\nabla}\cdot\left[\boldsymbol{v}f\right], \tag{A.16}$$

where D is the diffusion constant and $\boldsymbol{v}$ is the drift velocity. Consider the steady state problem, $\frac{\partial f}{\partial t} = 0$,

$$\begin{aligned} 0 &= D\nabla^2 f - \boldsymbol{\nabla} \cdot [\boldsymbol{v} f] \\ &= D\nabla^2 f - f\left(\boldsymbol{\nabla} \cdot \boldsymbol{v}\right) - \left(\boldsymbol{\nabla} f\right) \cdot \boldsymbol{v}. \end{aligned} \tag{A.17}$$

In order to cancel the Laplacian term, $\boldsymbol{v}$ must be of the form,

$$\boldsymbol{v} = g\boldsymbol{\nabla} f. \tag{A.18}$$

Thus, the steady state problem reduces to

$$\begin{aligned} 0 &= D\nabla^2 f - f\left(\boldsymbol{\nabla} \cdot [g\boldsymbol{\nabla} f]\right) - \left(\boldsymbol{\nabla} f\right) \cdot [g\boldsymbol{\nabla} f] \\ &= D\nabla^2 f - f\left(\boldsymbol{\nabla} g \cdot \boldsymbol{\nabla} f\right) - fg\nabla^2 f - \left(\boldsymbol{\nabla} f\right) \cdot [g\boldsymbol{\nabla} f]. \end{aligned} \tag{A.19}$$

To cancel that Laplacian term,

$$g = \frac{D}{f}. \tag{A.20}$$

In fact, for this choice of g the equation is satisfied,

$$\begin{aligned} 0 =& D\nabla^2 f - f\left[\boldsymbol{\nabla}\left(\frac{D}{f}\right) \cdot \boldsymbol{\nabla} f\right] - f\left(\frac{D}{f}\right)\nabla^2 f - \left(\boldsymbol{\nabla} f\right) \cdot \left[\left(\frac{D}{f}\right)\boldsymbol{\nabla} f\right] \\ =& D\nabla^2 f + \frac{D\boldsymbol{\nabla} f}{f} \cdot \boldsymbol{\nabla} f - D\nabla^2 f - \left(\boldsymbol{\nabla} f\right) \cdot \left[\left(\frac{D}{f}\right)\boldsymbol{\nabla} f\right] \\ =& 0. \end{aligned} \tag{A.21}$$

Supposing we had the Green's function for the Fokker-Plank PDE one could evolve an initial condition to the desired steady state distribution,

$$f(\boldsymbol{x}, t) = \int G(\boldsymbol{x}, \boldsymbol{x}', t) f(\boldsymbol{x}', 0) \tag{A.22}$$

However, the Green's function for the Fokker-Planck PDE is not known for a spatially varying drift velocity like we have here. If we limit ourselves to small time steps,

$$f(\boldsymbol{x}, \delta t) = \int G(\boldsymbol{x}, \boldsymbol{x}', \delta t) f(\boldsymbol{x}', 0), \tag{A.23}$$

where $\delta t \ll 1$, then the drift velocity can approximately be taken as a constant over one evolution of the distribution under the action of the Green's function. In that case, the Green's function is given by

$$G(\boldsymbol{x}, \boldsymbol{x}', \delta t) = \frac{e^{-\frac{(\boldsymbol{x}-\boldsymbol{x}'-\boldsymbol{v}(\boldsymbol{x}')\delta t)^2}{4Dt}}}{(4\pi Dt)^{3N/2}}. \tag{A.24}$$

Note that one could evolve an initial condition with the repeated application of the short time Green's function. However, the final distribution will only be the desired steady-state distribution in the $\delta t = 0$ limit (known as time-step error). A possible approach is to perform calculations with several time steps and extrapolate to $\delta t = 0$. This is time consuming and not necessary.

We can use the Metropolis-Hastings formalism that was developed earlier with

$$T_{\boldsymbol{x},\boldsymbol{x}'} = \frac{e^{-\frac{(\boldsymbol{x}-\boldsymbol{x}'-\boldsymbol{v}(\boldsymbol{x}')\delta t)^2}{4Dt}}}{(4\pi Dt)^{3N/2}}. \tag{A.25}$$

In conjunction with acceptance matrix for the Metropolis-Hastings method we are able to avoid the time step error. However, the quality of the proposal matrix does depend on δt.

A.2 Monte Carlo Estimators

A.2.1 Mean

Suppose that for a probability distribution, $\rho(\boldsymbol{R})$, we want to calculate the expectation value of an observable, $A(\boldsymbol{R})$. This is computed exactly as

$$\langle A \rangle_\rho = \int d\boldsymbol{R}\, \rho(\boldsymbol{R}) A(\boldsymbol{R}). \tag{A.26}$$

A Monte Carlo estimate, $\overline{A}$, of $\langle A \rangle_\rho$ is given by

$$\boxed{\overline{A} = \frac{1}{N} \sum_{i=1}^{N} A_i,} \tag{A.27}$$

where the set of N configurations, $\{\boldsymbol{R}_i | i = 1, N\}$, are chosen randomly (but not necessarily independently) with probability $\rho(\boldsymbol{R}_i)$, and $A_i \equiv A(\boldsymbol{R}_i)$. The estimator for the mean in Eqn. (A.27) is an unbiased estimator,

$$\begin{aligned}
\left\langle \overline{A} \right\rangle_\rho &= \left\langle \frac{1}{N} \sum_{i=1}^{N} A_i \right\rangle_\rho \\
\text{(Linearity of expection value)} \qquad &= \frac{1}{N} \sum_{i=1}^{N} \langle A_i \rangle_\rho \\
(\boldsymbol{R}_i \text{ sampled from } \rho(\boldsymbol{R}_i)) \qquad &= \frac{1}{N} \sum_{i=1}^{N} \langle A \rangle_\rho \\
&= \langle A \rangle_\rho
\end{aligned} \tag{A.28}$$

A.2.2 Statistical Error and Variance

The mean-square deviation of the estimator for the mean, $\overline{A}$, from the exact expecation value, $\langle A \rangle_\rho$ is

$$\begin{aligned}
\left\langle \left(\overline{A} - \langle A \rangle_\rho \right)^2 \right\rangle_\rho &= \left\langle \overline{A}^2 \right\rangle_\rho - \langle A \rangle_\rho^2 \\
&= \left\langle \frac{1}{N^2} \sum_{i,j=1}^{N} A_i A_j \right\rangle_\rho - \langle A \rangle_\rho^2 \\
&= \frac{1}{N} \left\langle A^2 \right\rangle_\rho - \langle A \rangle_\rho^2 + \frac{1}{N^2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \langle A_i A_j \rangle_\rho \\
&= \frac{1}{N} \underbrace{\left(\left\langle A^2 \right\rangle_\rho - \langle A \rangle_\rho^2 \right)}_{\text{Exact variance of A}} + \frac{1-N}{N} \langle A \rangle_\rho^2 + \frac{1}{N^2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \langle A_i A_j \rangle_\rho \\
&\equiv \frac{1}{N} \sigma_A^2 + \frac{1-N}{N} \langle A \rangle_\rho^2 + \frac{1}{N^2} \sum_{\substack{i,j=1 \\ i \neq j}}^{N} \langle A_i A_j \rangle_\rho \\
&= \frac{1}{N} \sigma_A^2 + \frac{2}{N^2} \sum_{\substack{i,j=1 \\ i < j}}^{N} \left[\langle A_i A_j \rangle_\rho - \langle A \rangle_\rho^2 \right] \\
&= \frac{1}{N} \sigma_A^2 + \frac{2}{N^2} \sum_{i=1}^{N} \sum_{t=1}^{N-i} \left[\langle A_i A_{i+t} \rangle_\rho - \langle A \rangle_\rho^2 \right], \qquad \text{(A.29)}
\end{aligned}$$

but the expectation values are over the same distribution, ρ, and the correlation between samples can only depend on the distance between them, t, so that each term in the sum over i is equivalent. Consequently, we set $i = 1$,

$$= \frac{1}{N} \sigma_A^2 + \frac{2}{N} \sum_{t=1}^{N-1} \left[\langle A_1 A_{1+t} \rangle_\rho - \langle A \rangle_\rho^2 \right]. \qquad \text{(A.30)}$$

Note that for independent samples $\boldsymbol{R}_i$, $\langle A_1 A_{1+t} \rangle_\rho = \langle A_1 \rangle_\rho \langle A_{1+t} \rangle_\rho = \langle A \rangle_\rho^2$. When there is some correlation between consecutive samples (for instance samples generated by a Markov process), $\langle A_1 A_{1+t} \rangle_\rho = \langle A \rangle_\rho^2$ does not generally hold. However, if N is large compared to the "time" it takes for correlations between

samples to fall off, i.e. $\exists\, t' \ll N \ni \langle A_1 A_{1+t}\rangle_\rho = \langle A\rangle_\rho^2 \ \forall\, t > t'$, then

$$\approx \frac{1}{N}\sigma_A^2 + \frac{2}{N}\sum_{t=1}^{\infty}\left[\langle A_1 A_{1+t}\rangle_\rho - \langle A\rangle_\rho^2\right]. \tag{A.31}$$

Now the integrated autocorrelation time of the observable A is defined as

$$\tau_A \equiv \frac{\sum_{t=1}^{\infty}\left[\langle A_1 A_{1+t}\rangle_\rho - \langle A\rangle_\rho^2\right]}{\sigma_A^2}, \tag{A.32}$$

so that

$$\left\langle \left(\overline{A} - \langle A\rangle_\rho\right)^2 \right\rangle_\rho = \frac{1}{N}\sigma_A^2(1+2\tau_A). \tag{A.33}$$

Since each $\boldsymbol{R}_i$ is a random variable sampled from $\rho(\boldsymbol{R})$, A_i is a random variable. If the variance of the distribution for A is finite then (need to verify this) the sum over the A_i values will satisfy the central limit theorem so that the root-mean-square deviation of the sample mean, $\overline{A}$, from the true expectation value, $\langle A\rangle_\rho$, can be interpreted as a statistical error

$$\Delta_A = \sqrt{\left\langle \left(\overline{A} - \langle A\rangle_\rho\right)^2 \right\rangle_\rho} \tag{A.34}$$

$$= \sqrt{\frac{1}{N}\sigma_A^2(1+2\tau_A)}. \tag{A.35}$$

In particular, $\langle A\rangle_\rho \in \{\overline{A}-\Delta_A, \overline{A}+\Delta_A\}$ 68% of the time, $\langle A\rangle_\rho \in \{\overline{A}-2\Delta_A, \overline{A}+2\Delta_A\}$ 95% of the time, etc. Since $\langle A\rangle_\rho$ is being estimated by $\overline{A}$ we need an estimate for Δ_A that does not involve $\langle A\rangle_\rho$. Consider the expectation value of the estimator $\overline{A^2} - \overline{A}^2$,

$$\begin{aligned}
\left\langle \overline{A^2} - \overline{A}^2 \right\rangle_\rho &= \left\langle \frac{1}{N}\sum_{i=1}^{N} A_i^2 - \frac{1}{N^2}\sum_{i,j=1}^{N} A_i A_j \right\rangle_\rho \\
&= \frac{N-1}{N}\left\langle A^2\right\rangle_\rho - \frac{1}{N^2}\sum_{\substack{i,j=1 \\ i\neq j}}^{N}\left\langle A_i A_j\right\rangle_\rho \\
&= \frac{N-1}{N}\sigma_A^2 - \frac{1}{N^2}\sum_{\substack{i,j=1 \\ i\neq j}}^{N}\left[\langle A_i A_j\rangle_\rho - \langle A\rangle_\rho^2\right] \\
&= \frac{N-(1+2\tau_A)}{N}\sigma_A^2.
\end{aligned} \tag{A.36}$$

The unbiased estimator for the variance is then

$$\boxed{\overline{\sigma_A^2} = \frac{1}{1 - \left(\frac{1+2\tau_A}{N}\right)} \left[\overline{A^2} - \overline{A}^2\right].} \tag{A.37}$$

Finally, from Eqn. (A.35) the estimator for the statistical error is

$$\boxed{\overline{\Delta_A} = \sqrt{\frac{1}{\left(\frac{N}{1+2\tau_A}\right) - 1} \left[\overline{A^2} - \overline{A}^2\right]}.} \tag{A.38}$$

Thus, whenever there is correlation between samples our effective number of samples is reduced by a factor of $1 + 2\tau_A$. The factor $1 + 2\tau_A$ is essentially how many Monte Carlo steps there are between independent measurements of the observable A. If $\tau_A = 0$ we recover the standard expression for the estimator of the statistical error

$$\overline{\Delta_A} = \sqrt{\frac{1}{N-1} \left[\overline{A^2} - \overline{A}^2\right]}. \tag{A.39}$$

From Eqn. (A.38) it is clear that τ_A is necessary to calculate the statistical error for the estimate of the mean, $\overline{A}$. However, calculating the autocorrelation time via Eqn. (A.32) is not practical. The autocorrelation time and correct statistical error bar can be computed with a binning analysis.

A.2.3 Covariance

The unbiased estimator of the covariance is (assuming independent samples)

$$\begin{aligned}
\overline{\mathrm{cov}(A,B)} &= \frac{1}{N-1} \sum_{i=1}^{N} \left(A_i - \overline{A}\right) \left(B_i - \overline{B}\right) \\
&= \frac{1}{N-1} \left[\sum_{i=1}^{N} A_i B_i - \frac{1}{N} \sum_{i,j=1}^{N} A_i B_j\right] \\
&= \frac{N}{N-1} \left[\overline{AB} - \overline{A}\,\overline{B}\right],
\end{aligned} \tag{A.40}$$

since

$$\begin{aligned}
\left\langle \overline{\mathrm{cov}(A,B)} \right\rangle_\rho &= \frac{1}{N-1}\left[\sum_{i=1}^{N}\langle A_iB_i\rangle_\rho - \frac{1}{N}\sum_{i,j=1}^{N}\langle A_iB_j\rangle_\rho\right] \\
&= \frac{1}{N-1}\left[\sum_{i=1}^{N}\langle A_iB_i\rangle_\rho - \frac{1}{N}\sum_{i=1}^{N}\langle A_iB_i\rangle_\rho - \frac{1}{N}\sum_{i\neq j}^{N}\langle A_iB_j\rangle_\rho\right] \\
&= \frac{1}{N-1}\left[\frac{N-1}{N}\sum_{i=1}^{N}\langle AB\rangle_\rho - \frac{1}{N}\sum_{i\neq j}^{N}\langle A\rangle_\rho\langle B\rangle_\rho\right] \\
&= \langle AB\rangle_\rho - \langle A\rangle_\rho\langle B\rangle_\rho \\
&= \mathrm{cov}(A,B). \qquad \text{(A.41)}
\end{aligned}$$

A.2.4 Differences of Operators

Suppose that we have an observable O that can be written as a difference of two operators

$$O = A - B. \qquad \text{(A.42)}$$

Clearly, the unbiased estimator for $\langle O\rangle_\rho$ is

$$\overline{O} = \overline{A} - \overline{B} \qquad \text{(A.43)}$$

by the linearity of the expectation value. Before calculating the statistical error, of our estimator we consider the quantity

$$\begin{aligned}
\left\langle \left[\overline{A} - \langle A\rangle_\rho\right]\left[\overline{B} - \langle B\rangle_\rho\right]\right\rangle_\rho &= \left\langle \overline{A}\,\overline{B} - \langle A\rangle_\rho\overline{B} - \overline{A}\langle B\rangle_\rho + \langle A\rangle_\rho\langle B\rangle_\rho\right\rangle_\rho \\
&= \left\langle \overline{A}\,\overline{B}\right\rangle_\rho - \langle A\rangle_\rho\langle B\rangle_\rho \\
&= \frac{1}{N^2}\left[\sum_{i=1}^{N}\langle A_iB_i\rangle_\rho + \sum_{i\neq j}^{N}\langle A_iB_j\rangle_\rho\right] - \langle A\rangle_\rho\langle B\rangle_\rho \\
\text{(independent samples)} \qquad &= \frac{1}{N^2}\left[N\langle AB\rangle_\rho + N(N-1)\langle A\rangle_\rho\langle B\rangle_\rho\right] - \langle A\rangle_\rho\langle B\rangle_\rho \\
&= \frac{1}{N}\mathrm{cov}(A,B). \qquad \text{(A.44)}
\end{aligned}$$

The exact statistical error given by Eqn. (A.34) is

$$\begin{aligned}
\Delta_O &= \sqrt{\left\langle \left(\overline{O} - \langle O \rangle_\rho\right)^2 \right\rangle_\rho} \\
&= \sqrt{\left\langle \left(\left[\overline{A} - \langle A \rangle_\rho\right] - \left[\overline{B} - \langle B \rangle_\rho\right] \right)^2 \right\rangle_\rho} \\
\text{(Eqns. (A.35), (A.44))} \qquad &= \sqrt{\Delta_A^2 + \Delta_B^2 - \frac{2}{N}\mathrm{cov}(A,B)}. \qquad \text{(A.45)}
\end{aligned}$$

Further, from Eqn. (A.38) we have the estimator for the statistical error

$$\begin{aligned}
\overline{\Delta_O} &= \sqrt{\frac{1}{N-1}\left[\overline{O^2} - \overline{O}^2\right]} \\
&= \sqrt{\frac{1}{N-1}\left[\overline{A^2 - 2AB + B^2} - \overline{A}^2 - \overline{B}^2 + 2\overline{A}\,\overline{B}\right]} \\
&= \sqrt{\overline{\Delta_A}^2 + \overline{\Delta_B}^2 - \frac{2}{N}\overline{\mathrm{cov}(A,B)}} \qquad \text{(A.46)}
\end{aligned}$$

Note that this is the result we would have obtained taking the exact error, Eqn. (A.45), and replacing the individual pieces with their estimators.

A.2.5 Products of Expectation Values

Suppose that we want to estimate the quantity $\langle A \rangle_\rho \langle B \rangle_\rho$. Naively one would estimate the quantity with $\overline{A}\,\overline{B}$, but this would be a biased estimator since

$$\begin{aligned}
\left\langle \overline{A}\,\overline{B} \right\rangle_\rho &= \left\langle \frac{1}{N^2}\sum_{i,j=1}^{N} A_i B_j \right\rangle_\rho \\
\text{(independent samples)} \qquad &= \frac{1}{N^2}\sum_{i=1}^{N} \langle AB \rangle_\rho + \frac{1}{N^2}\sum_{i\neq j}^{N} \langle A \rangle_\rho \langle B \rangle_\rho \\
&= \frac{1}{N}\langle AB \rangle_\rho + \frac{N-1}{N}\langle A \rangle_\rho \langle B \rangle_\rho \\
&= \langle A \rangle_\rho \langle B \rangle_\rho \left[1 + \frac{1}{N}\frac{\mathrm{cov}(A,B)}{\langle A \rangle_\rho \langle B \rangle_\rho}\right] \qquad \text{(A.47)}
\end{aligned}$$

Notice the bias is $\mathcal{O}(1/N)$. Note that it is clear how to calculate the unbiased estimator from the above. We simply need to remove the term involving $\langle AB \rangle_\rho$ and change the constant factor of the $\langle A \rangle_\rho \langle B \rangle_\rho$ term. Now, $\overline{AB}$ is the unbiased estimator of $\langle AB \rangle_\rho$ so that the unbiased estimator of the product of expectation values can be written as

$$\boxed{\overline{P} = \frac{N}{N-1}\overline{A}\,\overline{B} - \frac{1}{N-1}\overline{AB}}. \tag{A.48}$$

Alternatively, if we note that $\langle A \rangle_\rho \langle B \rangle_\rho = \langle AB \rangle_\rho - \text{cov}(A, B)$ and use the linearity of the expectation value we have

$$\begin{aligned} \overline{P} &= \overline{AB} - \overline{\text{cov}(A,B)} \\ \text{(Eqn. (A.40))} \qquad &= \overline{AB} - \frac{N}{N-1}\left[\overline{AB} - \overline{A}\,\overline{B}\right] \\ &= \frac{N}{N-1}\overline{A}\,\overline{B} - \frac{1}{N-1}\overline{AB} \end{aligned} \tag{A.49}$$

which is the same result.

A.3 Alternative Approach to Monte Carlo Estimators

In this section we provide an alternative method for determining the bias of an estimator. We first present the case of the product of two estimators which was done exactly earlier in the notes. Finally, we present the case of the ratio of two estimators which cannot be calculated exactly. For these two cases one must realize that our estimators can be written as

$$\overline{A} = \langle A \rangle_\rho + \eta_A \tag{A.50}$$

$$\overline{B} = \langle B \rangle_\rho + \eta_B \tag{A.51}$$

where η_X ($X = A, B$) is a random variable with mean 0 and variance $\frac{\sigma_X^2}{N}$

$$\langle \eta_X \rangle_\rho = 0 \tag{A.52}$$

$$\langle \eta_X^2 \rangle_\rho = \frac{\sigma_X^2}{N}. \tag{A.53}$$

A.3.1 Products of Expectation Values

Notice that Eqn. (A.40) implies

$$\begin{aligned}
\langle \overline{A}\,\overline{B} \rangle_\rho &= \langle \overline{AB} \rangle_\rho - \frac{N-1}{N}\mathrm{cov}(A,B) \\
&= \langle \overline{AB} \rangle_\rho - \langle A \rangle_\rho \langle B \rangle_\rho + \langle A \rangle_\rho \langle B \rangle_\rho - \frac{N-1}{N}\mathrm{cov}(A,B) \\
&= \langle A \rangle_\rho \langle B \rangle_\rho + \frac{\mathrm{cov}(A,B)}{N}.
\end{aligned} \tag{A.54}$$

Hence,

$$\langle \eta_a \eta_B \rangle_\rho = \frac{\mathrm{cov}(A,B)}{N}. \tag{A.55}$$

The naive guess for an estimator of the quantity $\langle A \rangle_\rho \langle B \rangle_\rho$ is $\overline{A}\,\overline{B}$. However, this estimator is biased

$$\begin{aligned}
\langle \overline{A}\,\overline{B} \rangle_\rho &= \Big\langle \Big(\langle A \rangle_\rho + \eta_A \Big) \Big(\langle B \rangle_\rho + \eta_B \Big) \Big\rangle_\rho \\
&= \Big\langle \langle A \rangle_\rho \langle B \rangle_\rho + \eta_A \langle B \rangle_\rho + \langle A \rangle_\rho \eta_B + \eta_A \eta_B \Big\rangle_\rho \\
&= \langle A \rangle_\rho \langle B \rangle_\rho + \langle \eta_A \rangle_\rho \langle B \rangle_\rho + \langle A \rangle_\rho \langle \eta_B \rangle_\rho + \langle \eta_A \eta_B \rangle_\rho \\
&= \langle A \rangle_\rho \langle B \rangle_\rho + \langle \eta_A \eta_B \rangle_\rho \\
&= \langle A \rangle_\rho \langle B \rangle_\rho + \frac{\mathrm{cov}(A,B)}{N} \\
&= \langle A \rangle_\rho \langle B \rangle_\rho \left[1 + \frac{1}{N} \frac{\mathrm{cov}(A,B)}{\langle A \rangle_\rho \langle B \rangle_\rho} \right],
\end{aligned} \tag{A.56}$$

which agrees with Eqn. (A.47). Note that if $\mathrm{cov}(A,B) = 0$, then the $\mathcal{O}(1/N)$ bias disappears.

We would like to calculate the variance of this estimator. Consider,

$$\begin{aligned}\left\langle \left(\overline{A}\,\overline{B}\right)^2 \right\rangle_\rho &= \left\langle \left(\langle A\rangle_\rho + \eta_A\right)^2 \left(\langle B\rangle_\rho + \eta_B\right)^2 \right\rangle_\rho \\ &= \left\langle \left(\langle A\rangle_\rho^2 + 2\langle A\rangle_\rho \eta_A + \eta_A^2\right)\left(\langle B\rangle_\rho^2 + 2\langle B\rangle_\rho \eta_B + \eta_B^2\right) \right\rangle_\rho \\ &= \langle A\rangle_\rho^2 \langle B\rangle_\rho^2 \left[1 + \frac{1}{N}\left(\frac{\sigma_B^2}{\langle B\rangle_\rho^2} + 4\frac{\mathrm{cov}(A,B)}{\langle A\rangle_\rho \langle B\rangle_\rho} + \frac{\sigma_A^2}{\langle A\rangle_\rho^2}\right)\right], \end{aligned} \tag{A.57}$$

where we have kept only terms of Hence, keeping terms of $\mathcal{O}(\frac{1}{N})$. Hence, to $\mathcal{O}(\frac{1}{N})$

$$\mathrm{Var}\left(\overline{A}\,\overline{B}\right) = \boxed{\frac{1}{N}\langle A\rangle_\rho^2 \langle B\rangle_\rho^2 \left(\frac{\sigma_A^2}{\langle A\rangle_\rho^2} + \frac{\sigma_B^2}{\langle B\rangle_\rho^2} + 2\frac{\mathrm{cov}(A,B)}{\langle A\rangle_\rho \langle B\rangle_\rho}\right)}. \tag{A.58}$$

As a final note, notice that to $\mathcal{O}(\eta)$, $\overline{A}\,\overline{B}$ is normally distributed

$$\begin{aligned}\overline{A}\,\overline{B} &= \langle A\rangle_\rho \langle B\rangle_\rho + \eta_A \langle B\rangle_\rho + \langle A\rangle_\rho \eta_B + \eta_A\eta_B \\ (\text{keep } \mathcal{O}(\eta)) \qquad &= \langle A\rangle_\rho \langle B\rangle_\rho \left[1 + \frac{\eta_A}{\langle A\rangle_\rho} + \frac{\eta_B}{\langle B\rangle_\rho}\right] \end{aligned} \tag{A.59}$$

So, to $\mathcal{O}(\eta)$,

$$\overline{A}\,\overline{B} \sim N\left(\langle A\rangle_\rho \langle B\rangle_\rho, \frac{1}{N}\langle A\rangle_\rho^2 \langle B\rangle_\rho^2 \left[\frac{\sigma_A^2}{\langle A\rangle_\rho^2} + \frac{\sigma_B^2}{\langle B\rangle_\rho^2} + 2\frac{\mathrm{cov}(A,B)}{\langle A\rangle_\rho \langle B\rangle_\rho}\right]\right). \tag{A.60}$$

However, to $\mathcal{O}(\eta^2)$, $\bar{A}\,\bar{B}$ involves a sum of normally distributed random variables and chi-squared random variables, so its distribution is quite complicated.

A.3.2 Ratios of Expectation Values

The naive guess for an estimator of the quantity $\frac{\langle A\rangle_\rho}{\langle B\rangle_\rho}$ is $\frac{\overline{A}}{\overline{B}}$. With the tools presented in this section we can now calculate the bias

$$\begin{aligned}
\left\langle \frac{\overline{A}}{\overline{B}} \right\rangle_\rho &= \left\langle \frac{\langle A\rangle_\rho + \eta_A}{\langle B\rangle_\rho + \eta_B} \right\rangle_\rho \\
&= \frac{1}{\langle B\rangle_\rho} \left\langle \frac{\langle A\rangle_\rho + \eta_A}{1 + \frac{\eta_B}{\langle B\rangle_\rho}} \right\rangle_\rho \\
(N \gg 1 \to \frac{\eta_B}{\langle B\rangle_\rho} \ll 1) \qquad &\approx \frac{1}{\langle B\rangle_\rho} \left\langle \left(\langle A\rangle_\rho + \eta_A\right) \left(1 - \frac{\eta_B}{\langle B\rangle_\rho} + \frac{\eta_B^2}{\langle B\rangle_\rho^2}\right) \right\rangle_\rho \\
(\text{keep } \mathcal{O}(\eta^2)) \qquad &= \frac{1}{\langle B\rangle_\rho} \left(\langle A\rangle_\rho - \frac{\langle \eta_A \eta_B\rangle_\rho}{\langle B\rangle_\rho} + \langle A\rangle_\rho \frac{\langle \eta_B^2\rangle_\rho}{\langle B\rangle_\rho^2} \right) \\
&= \boxed{\frac{\langle A\rangle_\rho}{\langle B\rangle_\rho} \left[1 + \frac{1}{N} \left(\frac{\sigma_B^2}{\langle B\rangle_\rho^2} - \frac{\mathrm{cov}(A,B)}{\langle A\rangle_\rho \langle B\rangle_\rho} \right) \right]}. \qquad \text{(A.61)}
\end{aligned}$$

Note that the $\mathcal{O}(1/N)$ bias is nonzero even if $\mathrm{cov}(A,B) = 0$!

We would like to calculate the variance of this estimator. Consider,

$$\begin{aligned}
\left\langle \left(\frac{\overline{A}}{\overline{B}}\right)^2 \right\rangle_\rho &= \left\langle \left(\frac{\langle A\rangle_\rho + \eta_A}{\langle B\rangle_\rho + \eta_B}\right)^2 \right\rangle_\rho \\
&= \frac{1}{\langle B\rangle_\rho^2} \left\langle \left(\frac{\langle A\rangle_\rho + \eta_A}{1 + \frac{\eta_B}{\langle B\rangle_\rho}}\right)^2 \right\rangle_\rho \\
(\text{keep } \mathcal{O}(\eta^2)) \qquad &= \boxed{\frac{\langle A\rangle_\rho^2}{\langle B\rangle_\rho^2} \left[1 + \frac{1}{N} \left(3\frac{\sigma_B^2}{\langle B\rangle_\rho^2} - 4\frac{\mathrm{cov}(A,B)}{\langle A\rangle_\rho \langle B\rangle_\rho} + \frac{\sigma_A^2}{\langle A\rangle_\rho^2} \right) \right]}. \qquad \text{(A.62)}
\end{aligned}$$

Hence, keeping terms of $\mathcal{O}(\frac{1}{N})$ we have

$$\mathrm{Var}\left(\frac{\overline{A}}{\overline{B}}\right) = \boxed{\frac{1}{N} \frac{\langle A\rangle_\rho^2}{\langle B\rangle_\rho^2} \left[\frac{\sigma_A^2}{\langle A\rangle_\rho^2} + \frac{\sigma_B^2}{\langle B\rangle_\rho^2} - 2\frac{\mathrm{cov}(A,B)}{\langle A\rangle_\rho \langle B\rangle_\rho} \right]}. \qquad \text{(A.63)}$$

As a final note, notice that to $\mathcal{O}(\eta)$, $\frac{\overline{A}}{\overline{B}}$ is normally distributed

$$\begin{aligned}
\frac{\overline{A}}{\overline{B}} &= \frac{1}{\langle B\rangle_\rho}\left[\langle A\rangle_\rho + \eta_A - \langle A\rangle_\rho \frac{\eta_B}{\langle B\rangle_\rho} - \eta_A\frac{\eta_B}{\langle B\rangle_\rho} + \langle A\rangle_\rho \frac{\eta_B^2}{\langle B\rangle_\rho^2}\right] \\
&= \frac{\langle A\rangle_\rho}{\langle B\rangle_\rho}\left[1 + \frac{\eta_A}{\langle A\rangle_\rho} - \frac{\eta_B}{\langle B\rangle_\rho} - \frac{\eta_A\eta_B}{\langle A\rangle_\rho\langle B\rangle_\rho} + \frac{\eta_B^2}{\langle B\rangle_\rho^2}\right] \\
\text{(keep } \mathcal{O}(\eta)) \qquad &= \frac{\langle A\rangle_\rho}{\langle B\rangle_\rho}\left[1 + \frac{\eta_A}{\langle A\rangle_\rho} - \frac{\eta_B}{\langle B\rangle_\rho}\right] \qquad \text{(A.64)}
\end{aligned}$$

So, to $\mathcal{O}(\eta)$,

$$\frac{\overline{A}}{\overline{B}} \sim N\left(\frac{\langle A\rangle_\rho}{\langle B\rangle_\rho}, \frac{1}{N}\frac{\langle A\rangle_\rho^2}{\langle B\rangle_\rho^2}\left[\frac{\sigma_A^2}{\langle A\rangle_\rho^2} + \frac{\sigma_B^2}{\langle B\rangle_\rho^2} - 2\frac{\mathrm{cov}(A,B)}{\langle A\rangle_\rho\langle B\rangle_\rho}\right]\right). \qquad \text{(A.65)}$$

However, to $\mathcal{O}(\eta^2)$, $\frac{\bar{A}}{\bar{B}}$ involves a sum of normally distributed random variables and chi-squared random variables, so its distribution is quite complicated. Finally, as can be seen in Eqn. (A.64) the non-normality of $\frac{\bar{A}}{\bar{B}}$ becomes most severe for small $\langle B\rangle_\rho$ since the chi-squared term can become large.

APPENDIX B

APPENDIX FOR CHAPTER 2

B.1 Scaling Relations

To derive the scaling relation between N_n^ζ and N_n^1, consider

$$\begin{aligned} 1 &= \left(N_n^\zeta\right)^2 \int_0^\infty dr\ r^2\ r^{2(n-1)} e^{-2\frac{(\zeta r)^2}{1+\zeta r}} \\ &= \frac{1}{\zeta^{2n+1}} \left(N_n^\zeta\right)^2 \int_0^\infty du\ u^{2n} e^{-2\frac{u^2}{1+u}} \\ &= \frac{1}{\zeta^{2n+1}} \left(\frac{N_n^\zeta}{N_n^1}\right)^2 . \end{aligned} \tag{B.1}$$

Hence, the scaling relation for the normalization factor is

$$N_n^\zeta = \zeta^{n+1/2}\ N_n^1. \tag{B.2}$$

The values of N_n^1 are given in Table B.1.

Table B.1: Normalization factors for Gauss-Slater basis functions with unit exponent and principal quantum number n.

n	N_n^1
1	1.126467421
2	0.576609950
3	0.196581141
4	0.050275655
5	0.010280772

To derive the scaling relations for the parameters α_i^ζ and c_i^ζ in the Gaussian expansion of the Gauss-Slater functions, suppose the best-fit expansion for $\zeta = 1$

is

$$N_n^1 \, e^{-\frac{r^2}{1+r}} = \sum_i c_i^1 \sqrt{\frac{2(2\alpha_i^1)^{n+\frac{1}{2}}}{\Gamma(n+\frac{1}{2})}} \, e^{-\alpha_i^1 r^2}. \tag{B.3}$$

Using Eqn. (B.2) and performing the substitution $r \to \zeta r$ results in

$$\begin{aligned} N_n^\zeta \, e^{-\frac{(\zeta r)^2}{1+\zeta r}} &= \zeta^{n+1/2} \sum_i c_i^1 \sqrt{\frac{2(2\alpha_i^1)^{n+\frac{1}{2}}}{\Gamma(n+\frac{1}{2})}} \, e^{-\alpha_i^1 \zeta^2 r^2} \\ &= \sum_i c_i^1 \sqrt{\frac{2(2\alpha_i^1 \zeta^2)^{n+\frac{1}{2}}}{\Gamma(n+\frac{1}{2})}} \, e^{-\alpha_i^1 \zeta^2 r^2} \\ &= \sum_i c_i^\zeta \sqrt{\frac{2(2\alpha_i^\zeta)^{n+\frac{1}{2}}}{\Gamma(n+\frac{1}{2})}} \, e^{-\alpha_i^\zeta r^2}, \end{aligned} \tag{B.4}$$

where

$$\alpha_i^\zeta = \zeta^2 \alpha_i^1 \tag{B.5}$$

$$c_i^\zeta = c_i^1. \tag{B.6}$$

B.2 Spatial Derivatives

A general unnormalized radial basis function has the form

$$R_n^\zeta(r) = r^{n-1} e^{g^\zeta(r)}, \tag{B.7}$$

where $g^\zeta(r)$ is an arbitrary function. The gradient is

$$\boldsymbol{\nabla} R_n^\zeta(r) = \frac{\partial R_n^\zeta(r)}{\partial r} \hat{\boldsymbol{r}}, \tag{B.8}$$

where

$$\frac{\partial R_n^\zeta(r)}{\partial r} = R_n^\zeta(r) \left[\frac{(n-1)}{r} + \frac{\partial g^\zeta(r)}{\partial r} \right]. \tag{B.9}$$

The Laplacian is

$$\nabla^2 R_n^\zeta(r) = \frac{\partial^2 R_n^\zeta(r)}{\partial r^2} + \frac{2}{r} \frac{\partial R_n^\zeta(r)}{\partial r}, \tag{B.10}$$

where

$$\frac{\partial^2 R_n^\zeta(r)}{\partial r^2} = R_n^\zeta(r)\left[\frac{\partial^2 g^\zeta(r)}{\partial r^2} - \frac{(n-1)}{r^2}\right] + \frac{1}{R_n^\zeta(r)}\left(\frac{\partial R_n^\zeta(r)}{\partial r}\right)^2. \tag{B.11}$$

For Gauss-Slater functions,

$$g^\zeta(r) = -\frac{(\zeta r)^2}{1+\zeta r} \tag{B.12}$$

$$\frac{\partial g^\zeta(r)}{\partial r} = -\frac{r\zeta^2(2+\zeta r)}{(1+\zeta r)^2} \tag{B.13}$$

$$\frac{\partial^2 g^\zeta(r)}{\partial r^2} = -\frac{2\zeta^2}{(1+\zeta r)^3}. \tag{B.14}$$

For Gaussian functions,

$$g^\zeta(r) = -\zeta r^2 \tag{B.15}$$

$$\frac{\partial g^\zeta(r)}{\partial r} = -2\zeta r \tag{B.16}$$

$$\frac{\partial^2 g^\zeta(r)}{\partial r^2} = -2\zeta. \tag{B.17}$$

For Slater functions,

$$g^\zeta(r) = -\zeta r \tag{B.18}$$

$$\frac{\partial g^\zeta(r)}{\partial r} = -\zeta \tag{B.19}$$

$$\frac{\partial^2 g^\zeta(r)}{\partial r^2} = 0. \tag{B.20}$$

B.3 Parameter Derivatives

Wavefunction optimization via the linear method requires both the derivatives of the wavefunction with respect to the exponent parameters ζ, and the Hamiltonian acting on those derivatives. From Eqn. (B.2), the derivative of the normalization with respect to the exponent is

$$\frac{\partial N_n^\zeta}{\partial \zeta} = \frac{(n+1/2)}{\zeta} N_n^\zeta. \tag{B.21}$$

Now consider a general unnormalized radial basis function of the form

$$R_n^\zeta(r) = r^{n-1} e^{g^\zeta(r)}, \tag{B.22}$$

where $g^\zeta(r)$ is an arbitrary function. The derivative of the radial part of the wavefunction with respect to the exponent is

$$\frac{\partial R_n^\zeta(r)}{\partial \zeta} = f^\zeta(r) R_n^\zeta(r), \tag{B.23}$$

where

$$f^\zeta(r) \equiv \frac{\partial g^\zeta(r)}{\partial \zeta}. \tag{B.24}$$

The gradient is

$$\boldsymbol{\nabla}\left[\frac{\partial R_n^\zeta(r)}{\partial \zeta}\right] = \frac{\partial f^\zeta(r)}{\partial r} R_n^\zeta(r)\hat{\boldsymbol{r}} + f^\zeta(r)\left[\boldsymbol{\nabla} R_n^\zeta(r)\right]. \tag{B.25}$$

The Laplacian is

$$\nabla^2\left[\frac{\partial R_n^\zeta(r)}{\partial \zeta}\right] = \frac{2}{r}\frac{\partial f^\zeta(r)}{\partial r}\left[R_n^\zeta(r) + r\frac{\partial R_n^\zeta(r)}{\partial r}\right] + \frac{\partial^2 f^\zeta(r)}{\partial r^2} R_n^\zeta(r) + f^\zeta(r)\nabla^2 R_n^\zeta(r). \tag{B.26}$$

For Gauss-Slater functions,

$$f^\zeta(r) = -\frac{\zeta r^2(2+\zeta r)}{(1+\zeta r)^2} \tag{B.27}$$

$$\frac{\partial f^\zeta(r)}{\partial r} = -\frac{\zeta r[4+\zeta r(3+\zeta r)]}{(1+\zeta r)^3} \tag{B.28}$$

$$\frac{\partial^2 f^\zeta(r)}{\partial r^2} = \frac{2\zeta(\zeta r - 2)}{(1+\zeta r)^4}. \tag{B.29}$$

For Gaussian functions,

$$f^\zeta(r) = -r^2 \tag{B.30}$$

$$\frac{\partial f^\zeta(r)}{\partial r} = -2r \tag{B.31}$$

$$\frac{\partial^2 f^\zeta(r)}{\partial r^2} = -2. \tag{B.32}$$

For Slater functions,

$$f^{\zeta}(r) = -r \tag{B.33}$$

$$\frac{\partial f^{\zeta}(r)}{\partial r} = -1 \tag{B.34}$$

$$\frac{\partial^2 f^{\zeta}(r)}{\partial r^2} = 0. \tag{B.35}$$

B.4 Exponents

To promote use of this basis, exponents for each system studied are given in this appendix. Only exponents of the primitives are given, as the contracted functions are presented elsewhere [18]. Exponents for the ground state of carbon using CAS(4,13) CSFs, and, the ground state of carbon dimer using a single CSF are shown in Table B.2. Exponents for the lowest lying excited states of carbon with ${}^5S^o$, ${}^3P^o$, ${}^1D^o$, ${}^3F^o$ symmetries are shown in Table B.3. Exponents for the ground state of naphthalene are shown in Table B.4. In naphthalene, since atoms of the same atomic species are located at inequivalent geometrical locations, one could independently optimize the exponents for each inequivalent atom, but we have not done so because we expect the resulting gain to be small.

We have found that the carbon S and P exponents change relatively little from one molecule to another (though they do differ more for the atom) while there is considerable leeway in the D exponents (they change considerably even from one optimization to another for a given molecule). This is because the energy and σ are not as sensitive to the D basis functions as they do not appear in the ground-state determinant of the carbon atom. Hence it is possible to find an approximately optimal set of exponents for the atoms in a large molecule, by optimizing them for a small molecule with the same atoms.

Table B.2: Basis exponents for CAS(4,13) ground state of carbon and ground state of carbon dimer using G and GS basis functions. For each n, the nz basis includes $n-1$ S primitives, $n-1$ P primitives, $n-1$ D primitives, and $n-2$ F primitives (where applicable).

Type	Size	L	C Exp.	C_2 Exp.
G	$2z$	S	0.087	0.145
		P	0.129	0.196
		D	0.470	0.679
	$3z$	S	0.102	0.127
		S	0.676	0.998
		P	0.104	0.121
		P	0.270	0.423
		D	0.314	0.386
		D	0.982	1.099
		F	N/A	0.783
GS	$2z$	S	0.586	0.853
		P	0.984	1.162
		D	1.810	3.774
	$3z$	S	1.000	1.127
		S	1.258	1.570
		P	1.059	0.703
		P	1.750	1.416
		D	1.132	2.225
		D	1.981	3.228
		F	N/A	2.588

Table B.3: Basis exponents for the lowest lying excited states of carbon with ${}^5S^o$, ${}^3P^o$, ${}^1D^o$, ${}^3F^o$ symmetries. For each n, the nz basis includes $n-1$ S primitives, $n-1$ P primitives, and $n-1$ D primitives (where applicable).

Type	Size	L	${}^5S^o$ Exp.	${}^3P^o$ Exp.	${}^1D^o$ Exp.	${}^3F^o$ Exp.
G	$2z$	S	0.110	0.006	0.087	0.112
		P	0.150	0.157	0.109	0.766
		D	N/A	N/A	0.006	0.006
	$3z$	S	0.356	0.010	0.094	0.096
		S	2.145	0.284	0.496	0.689
		P	0.809	4.868	0.079	0.117
		P	2.262	6.707	0.854	0.524
		D	N/A	N/A	0.007	0.007
		D	N/A	N/A	0.407	0.822
GS	$2z$	S	1.960	0.059	0.756	0.570
		P	1.860	3.002	0.763	0.588
		D	N/A	N/A	0.351	0.293
	$3z$	S	1.148	0.371	0.160	0.894
		S	1.451	0.376	0.931	2.045
		P	1.137	0.691	0.043	1.326
		P	1.325	2.087	0.444	3.016
		D	N/A	N/A	0.238	0.295
		D	N/A	N/A	0.953	1.545

Table B.4: Basis exponents for ground state of naphthalene, $C_{10}H_8$, using G and GS basis functions. For carbon, the $2z$ basis includes 1 S primitive, 1 P primitive, 1 D primitive. For hydrogen, the $2z$ basis includes 1 S primitive.

Type	Size	L	$C_{10}H_8$ C Exp.	$C_{10}H_8$ H Exp.
G	$2z$	S	0.139	0.099
		P	0.191	N/A
		D	0.754	N/A
GS	$2z$	S	0.875	0.798
		P	1.118	N/A
		D	2.109	N/A

APPENDIX C

APPENDIX FOR CHAPTER 3

C.1 Contracted Functions

The contracted functions for the ANO-GS/GSn basis are given in Table C.1.

Table C.1: Contracted functions for the ANO-GS/GSn basis (continued on subsequent pages).

System	Ang. Mom.	Exp.	Coef.
H	S	0.04070137	0.01743456
		0.07733261	0.13755618
		0.14693196	0.28886067
		0.27917073	0.28678168
		0.53042439	0.19955827
		1.00780633	0.10852968
		1.91483203	0.05910120
		3.63818086	0.03532391
		6.91254363	0.00208123
He	S	0.10248700	0.04089897
		0.20497400	0.17797219
		0.40994800	0.28528092
		0.81989600	0.26264576
		1.63979200	0.19034914
		3.27958400	0.11079007
		6.55916800	0.05972012
		13.11833600	0.03700733

Table C.1 – continued

System	Ang. Mom.	Exp.	Coef.
		26.23667200	0.00265480
Li	S	0.00974359	0.00352139
		0.01851281	0.12903815
		0.03517434	0.44387349
		0.06683125	0.41065480
		0.12697938	0.14370785
		0.24126082	-0.03788394
		0.45839556	-0.07998393
		0.87095157	-0.08019704
		1.65480799	0.02185181
Be	S	0.02562891	-0.02641643
		0.05125781	-0.26469665
		0.10251562	-0.47655474
		0.20503125	-0.31422835
		0.41006250	-0.05743711
		0.82012500	0.08548245
		1.64025000	0.10710350
		3.28050000	-0.01780919
		6.56100000	0.00028805
	P	0.05062500	-0.08962232
		0.10631250	-0.39006883
		0.22325625	-0.41374468
		0.46883813	-0.16909999
		0.98456006	-0.04840263

Table C.1 – continued

System	Ang. Mom.	Exp.	Coef.
		2.06757613	-0.03205887
		4.34190988	-0.01041324
		9.11801074	0.00084677
		19.14782255	-0.00015624
B	S	0.05062500	0.06617715
		0.09618750	0.27188133
		0.18275625	0.40697090
		0.34723687	0.30816297
		0.65975006	0.09364792
		1.25352512	-0.05802092
		2.38169773	-0.11533259
		4.52522568	0.01237330
		8.59792879	0.00199977
	P	0.02560000	0.01149541
		0.05376000	0.12869896
		0.11289600	0.30526053
		0.23708160	0.33323434
		0.49787136	0.24570299
		1.04552986	0.12565772
		2.19561270	0.05791787
		4.61078666	0.01337759
		9.68265200	0.00106468
C	S	0.05062500	0.01598887
		0.10125000	0.16450719

Table C.1 – continued

System	Ang. Mom.	Exp.	Coef.
		0.20250000	0.38233257
		0.40500000	0.38771387
		0.81000000	0.19273788
		1.62000000	-0.02854048
		3.24000000	-0.12573272
		6.48000000	0.01468002
		12.96000000	-0.00056068
	P	0.05062500	0.03428372
		0.10125000	0.15793083
		0.20250000	0.28424988
		0.40500000	0.30233996
		0.81000000	0.23261465
		1.62000000	0.13296708
		3.24000000	0.06080434
		6.48000000	0.02999658
		12.96000000	0.00135166
N	S	0.07593750	-0.02686191
		0.15187500	-0.19321557
		0.30375000	-0.38474373
		0.60750000	-0.36536725
		1.21500000	-0.18207788
		2.43000000	0.04616424
		4.86000000	0.11528896
		9.72000000	-0.01986060

Table C.1 – continued

System	Ang. Mom.	Exp.	Coef.
		19.44000000	0.00078409
	P	0.06750000	0.02961387
		0.13500000	0.14041850
		0.27000000	0.26447192
		0.54000000	0.30063044
		1.08000000	0.24703829
		2.16000000	0.15132427
		4.32000000	0.07150723
		8.64000000	0.03478525
		17.28000000	0.00432024
O	S	0.11000000	0.03202827
		0.19800000	0.15307475
		0.35640000	0.29474179
		0.64152000	0.34704224
		1.15473600	0.23419103
		2.07852480	0.10026844
		3.74134464	-0.07861470
		6.73442035	-0.08350949
		12.12195663	0.01763657
	P	0.07700000	0.04130711
		0.15400000	0.14327885
		0.30800000	0.24460946
		0.61600000	0.27683163
		1.23200000	0.24908259

Table C.1 – continued

System	Ang. Mom.	Exp.	Coef.
		2.46400000	0.17441826
		4.92800000	0.08867451
		9.85600000	0.04488804
		19.71200000	0.00673237
F	S	0.13641020	0.02797138
		0.24553835	0.13686497
		0.44196904	0.29018339
		0.79554427	0.33878915
		1.43197968	0.25867176
		2.57756343	0.10284827
		4.63961418	-0.05605286
		8.35130552	-0.09992039
		15.03234993	0.01670889
	P	0.10248700	0.04254657
		0.20497400	0.13947990
		0.40994800	0.24175240
		0.81989600	0.27636193
		1.63979200	0.25237549
		3.27958400	0.17542641
		6.55916800	0.09152650
		13.11833600	0.04443016
		26.23667200	0.00513978
Ne	S	0.17569200	0.03422377
		0.31624560	0.14694149

Table C.1 – continued

System	Ang. Mom.	Exp.	Coef.
		0.56924208	0.29359528
		1.02463574	0.33571250
		1.84434434	0.24448271
		3.31981981	0.09720106
		5.97567566	-0.06977995
		10.75621619	-0.06483557
		19.36118914	0.00842959
	P	0.10717944	0.02332263
		0.21435888	0.10697547
		0.42871776	0.21486030
		0.85743552	0.27023961
		1.71487105	0.26653350
		3.42974210	0.20120113
		6.85948419	0.11703442
		13.71896838	0.06254500
		27.43793677	0.00658784
Na	S	0.01726136	0.15962472
		0.03279658	0.46763739
		0.06231350	0.41453230
		0.11839564	0.10369471
		0.22495172	-0.06422858
		0.42740827	-0.09822361
		0.81207570	-0.08827260
		1.54294384	0.03021669

Table C.1 – continued

System	Ang. Mom.	Exp.	Coef.
		2.93159329	-0.00088115
Mg	S	0.02297486	-0.05834372
		0.04365224	-0.34313617
		0.08293926	-0.48833841
		0.15758460	-0.26608755
		0.29941074	0.01344630
		0.56888040	0.13225689
		1.08087276	0.14219414
		2.05365825	-0.05439608
		3.90195067	0.00405301
	P	0.03057955	0.02695661
		0.05198523	0.17682202
		0.08837489	0.38169561
		0.15023731	0.37158738
		0.25540342	0.14842290
		0.43418581	-0.00124050
		0.73811588	-0.05271299
		1.25479700	-0.04013899
		2.13315490	0.01107743
Al	S	0.03700125	-0.07202255
		0.06660225	-0.28464829
		0.11988405	-0.45758868
		0.21579129	-0.31670557
		0.38842432	-0.06157971

Table C.1 – continued

System	Ang. Mom.	Exp.	Coef.
		0.69916377	0.14261991
		1.25849479	0.17701513
		2.26529063	-0.05806300
		4.07752313	0.00616560
	P	0.02779959	0.10096225
		0.05281921	0.27373295
		0.10035651	0.37457865
		0.19067736	0.27468527
		0.36228699	0.12036825
		0.68834529	-0.00528008
		1.30785604	-0.04982463
		2.48492648	0.00733696
		4.72136031	-0.00113882
Si	S	0.03240000	-0.00874843
		0.06156000	-0.13111530
		0.11696400	-0.41542276
		0.22223160	-0.45858890
		0.42224004	-0.21179781
		0.80225608	0.13567942
		1.52428654	0.21148601
		2.89614443	-0.04725110
		5.50267443	0.00229557
	P	0.03240000	-0.04259816
		0.06156000	-0.19674204

Table C.1 – continued

System	Ang. Mom.	Exp.	Coef.
		0.11696400	-0.35101560
		0.22223160	-0.33929651
		0.42224004	-0.19292528
		0.80225608	-0.02815469
		1.52428654	0.05474493
		2.89614443	-0.00309097
		5.50267443	0.00067220
P	S	0.05187485	-0.02375243
		0.09337473	-0.17887230
		0.16807451	-0.39775410
		0.30253412	-0.42608111
		0.54456142	-0.19606826
		0.98021055	0.09284981
		1.76437899	0.24630885
		3.17588218	-0.03921506
		5.71658793	-0.00215415
	P	0.04594973	0.04535552
		0.07811454	0.13096467
		0.13279472	0.26611330
		0.22575102	0.29765470
		0.38377674	0.25786675
		0.65242046	0.13484429
		1.10911478	0.02027525
		1.88549512	-0.05247665

Table C.1 – continued

System	Ang. Mom.	Exp.	Coef.
		3.20534170	-0.00416241
S	S	0.06152788	-0.01886366
		0.10459739	-0.11497881
		0.17781556	-0.31264645
		0.30228646	-0.39617586
		0.51388697	-0.31902178
		0.87360786	-0.06793635
		1.48513336	0.18081152
		2.52472671	0.18001334
		4.29203540	-0.05451818
	P	0.05350250	-0.06374846
		0.09630450	-0.16240617
		0.17334810	-0.28676895
		0.31202659	-0.30277046
		0.56164786	-0.24696068
		1.01096614	-0.11183221
		1.81973905	0.03134041
		3.27553029	0.02960453
		5.89595452	-0.00246437
Cl	S	0.09257500	-0.05241911
		0.16663500	-0.21568577
		0.29994300	-0.44405450
		0.53989740	-0.38395406
		0.97181532	-0.17146038

Table C.1 – continued

System	Ang. Mom.	Exp.	Coef.
		1.74926758	0.22058963
		3.14868164	0.20298184
		5.66762695	-0.07946711
		10.20172850	0.00689586
	P	0.06118046	-0.04168476
		0.11012482	-0.14017625
		0.19822468	-0.26402369
		0.35680443	-0.31550738
		0.64224797	-0.26657997
		1.15604634	-0.14441887
		2.08088342	0.02261610
		3.74559015	0.04758516
		6.74206227	-0.00681694
Ar	S	0.10000000	-0.03266327
		0.18000000	-0.18332457
		0.32400000	-0.40217886
		0.58320000	-0.43059024
		1.04976000	-0.21192824
		1.88956800	0.14966838
		3.40122240	0.27291260
		6.12220032	-0.08305932
		11.01996058	0.00413157
	P	0.06939182	-0.03009514
		0.12490528	-0.11975400

Table C.1 – continued

System	Ang. Mom.	Exp.	Coef.
		0.22482951	-0.24832306
		0.40469311	-0.31388823
		0.72844760	-0.28705129
		1.31120568	-0.16471762
		2.36017023	0.00637195
		4.24830641	0.05891777
		7.64695153	-0.00725650

C.2 GS Functions

C.2.1 Normalization of GS functions

The normalization factors for the radial part of the GS basis functions with unit exponent are given in Table C.2.

Table C.2: Normalization factors for GS basis functions with unit exponent and principal quantum number n.

n	N_n^1
1	1.126467421
2	0.576609950
3	0.196581141
4	0.050275655
5	0.010280772

C.2.2 Gaussian fits of GS functions

The expansions of the GS functions with unit exponent in a single Gaussian are given in Table C.3. The expansions of the GS functions with unit exponent in six Gaussians are given in Table C.4.

Table C.3: The expansions of the GS functions with unit exponent in a single Gaussian.

Function	Exp.	Coef.
GS-1S	0.20708437	1.00000000
GS-3S	0.04833286	1.00000000
GS-2P	0.15168443	1.00000000
GS-3P	0.08304123	1.00000000
GS-4P	0.05235243	1.00000000
GS-3D	0.11860705	1.00000000
GS-4D	0.07451191	1.00000000
GS-4F	0.09694663	1.00000000

Table C.4: The expansions of the GS functions with unit exponent in six Gaussians (continued on subsequent pages).

Function	Exp.	Coef.
GS-1S	0.06689139	0.18918908
	0.16047444	0.48208792
	0.39737459	0.32148998
	1.07508975	0.09402732
	3.46365258	0.01304374
	17.18649824	0.00063827

Table C.4 – continued

Function	Exp.	Coef.
GS-3S	0.02615360	0.10206590
	0.05465914	0.81024134
	0.11465363	0.31331649
	0.24039781	-0.22526824
	0.49848703	-0.08108560
	1.06677389	-0.03103457
GS-2P	0.05676262	0.20261442
	0.12307372	0.48590726
	0.26955306	0.32471268
	0.62972336	0.09930363
	1.67428436	0.01515182
	5.98403593	0.00089381
GS-3P	0.03085706	0.08339441
	0.06586273	0.57673030
	0.14128985	0.39763641
	0.28846603	0.01475027
	0.61860693	-0.01504486
	1.44581552	-0.01698559
GS-4P	0.02525014	0.09557434
	0.04880620	0.65897768
	0.09457222	0.36633078
	0.18377564	-0.08042587
	0.35102086	-0.04341420
	0.69569372	-0.01916470

Table C.4 – continued

Function	Exp.	Coef.
GS-3D	0.05169545	0.24526750
	0.10571780	0.49873695
	0.21472412	0.29991138
	0.45696287	0.08654449
	1.08040987	0.01296466
	3.24061913	0.00078847
GS-4D	0.03709411	0.22140826
	0.07022210	0.55582268
	0.13371057	0.28014977
	0.24564044	0.03219304
	0.40027045	-0.00451244
	0.96625100	-0.00763014
GS-4F	0.04845121	0.30307423
	0.09536478	0.50746119
	0.18409468	0.26328334
	0.36676059	0.06853641
	0.79713347	0.00963494
	2.12180291	0.00057001

C.2.3 Exponents

The exponents for the $2z$ ANO-GS basis are given in Table C.5. The exponents for the $2z$ ANO-GSn basis of lithium and sodium are given in Table C.6.

These are the only elements where the $2z$ ANO-GSn basis differs from the $2z$ ANO-GS basis. The exponents for the $3z$ ANO-GS basis are given in Table C.7. The exponents for the $3z$ ANO-GSn basis are given in Table C.8.

Table C.5: Exponents for $2z$ ANO-GS basis (continued on subsequent pages).

System	Function	Exp.
H	GS-1S	1.8008
	GS-2P	2.0063
He	GS-1S	3.0477
	GS-2P	2.7297
Li	GS-1S	0.4713
	GS-2P	0.7183
	GS-2P	1.9282
	GS-3D	0.8379
Be	GS-1S	0.4940
	GS-2P	0.6767
	GS-3D	1.2406
B	GS-1S	0.6835
	GS-2P	0.8468
	GS-3D	1.7179
C	GS-1S	0.8335
	GS-2P	1.0750
	GS-3D	2.2359
N	GS-1S	0.8903
	GS-2P	1.2640
	GS-3D	2.5937
O	GS-1S	0.9835

Table C.5 – continued

System	Function	Exp.
	GS-2P	1.2960
	GS-3D	2.8726
F	GS-1S	1.0734
	GS-2P	1.5617
	GS-3D	3.3773
Ne	GS-1S	1.2705
	GS-2P	1.8062
	GS-3D	4.1117
Na	GS-1S	0.4594
	GS-2P	0.7359
	GS-2P	1.1328
	GS-3D	0.7719
Mg	GS-1S	0.3548
	GS-2P	0.5323
	GS-3D	1.0473
Al	GS-1S	0.4798
	GS-2P	0.6281
	GS-3D	1.2257
Si	GS-1S	0.5335
	GS-2P	0.7775
	GS-3D	1.4937
P	GS-1S	0.6320
	GS-2P	0.9073
	GS-3D	1.7950

Table C.5 – continued

System	Function	Exp.
S	GS-1S	0.6656
	GS-2P	0.9748
	GS-3D	1.9992
Cl	GS-1S	0.6882
	GS-2P	1.0867
	GS-3D	2.1929
Ar	GS-1S	0.6914
	GS-2P	1.2242
	GS-3D	2.4625

Table C.6: Exponents for $2z$ ANO-GSn bases of lithium and sodium.

System	Function	Exp.
Li	GS-1S	0.4719
	GS-2P	1.5344
	GS-3P	0.9062
	GS-3D	0.8438
Na	GS-1S	0.4594
	GS-2P	1.0875
	GS-3P	0.8750
	GS-3D	0.7750

Table C.7: Exponents for $3z$ ANO-GS basis (continued on subsequent pages).

System	Function	Exp.
H	GS-1S	1.0180
	GS-1S	1.3266
	GS-2P	1.8750
	GS-2P	2.5328
	GS-3D	2.7219
He	GS-1S	1.5133
	GS-1S	2.4992
	GS-2P	2.5180
	GS-2P	4.0086
	GS-3D	3.8367
Li	GS-1S	0.5938
	GS-1S	0.6000
	GS-2P	0.8031
	GS-2P	1.1641
	GS-2P	1.9031
	GS-3D	0.7625
	GS-3D	0.7750
	GS-4F	1.0875
Be	GS-1S	0.8344
	GS-1S	0.8516
	GS-2P	0.7875
	GS-2P	2.2375
	GS-3D	1.0906
	GS-3D	1.0938

Table C.7 – continued

System	Function	Exp.
	GS-4F	1.5781
B	GS-1S	0.9961
	GS-1S	1.2078
	GS-2P	0.9828
	GS-2P	1.3445
	GS-3D	1.4289
	GS-3D	1.8094
	GS-4F	2.2172
C	GS-1S	1.1414
	GS-1S	1.6055
	GS-2P	1.2047
	GS-2P	1.6383
	GS-3D	1.9766
	GS-3D	2.4367
	GS-4F	2.7523
N	GS-1S	1.2094
	GS-1S	1.8875
	GS-2P	1.3188
	GS-2P	3.7000
	GS-3D	2.2062
	GS-3D	3.0000
	GS-4F	3.2437
O	GS-1S	1.3344
	GS-1S	2.3312

Table C.7 – continued

System	Function	Exp.
	GS-2P	1.3219
	GS-2P	4.0687
	GS-3D	2.4906
	GS-3D	3.6063
	GS-4F	3.5219
F	GS-1S	1.7125
	GS-1S	2.5938
	GS-2P	1.6531
	GS-2P	2.4000
	GS-3D	2.7812
	GS-3D	4.2594
	GS-4F	4.0438
Ne	GS-1S	2.1031
	GS-1S	2.8125
	GS-2P	1.8844
	GS-2P	3.2000
	GS-3D	3.2656
	GS-3D	4.9531
	GS-4F	4.8812
Na	GS-1S	0.5406
	GS-1S	0.5719
	GS-2P	0.5813
	GS-2P	0.9031
	GS-2P	0.9187

Table C.7 – continued

System	Function	Exp.
	GS-3D	0.6687
	GS-3D	0.6656
	GS-4F	1.0531
Mg	GS-1S	0.7063
	GS-1S	0.7125
	GS-2P	1.0344
	GS-2P	1.0469
	GS-3D	1.1477
	GS-3D	1.1586
	GS-4F	1.2180
Al	GS-1S	0.8500
	GS-1S	0.8656
	GS-2P	0.7188
	GS-2P	1.4000
	GS-3D	1.0437
	GS-3D	1.2594
	GS-4F	1.4656
Si	GS-1S	0.9281
	GS-1S	0.9500
	GS-2P	0.8656
	GS-2P	1.7219
	GS-3D	1.1812
	GS-3D	1.6187
	GS-4F	1.7875

Table C.7 – continued

System	Function	Exp.
P	GS-1S	1.0844
	GS-1S	1.1125
	GS-2P	1.0250
	GS-2P	2.0312
	GS-3D	1.4781
	GS-3D	1.9625
	GS-4F	2.0781
S	GS-1S	1.2125
	GS-1S	1.2281
	GS-2P	1.0125
	GS-2P	2.2687
	GS-3D	1.5750
	GS-3D	2.2000
	GS-4F	2.2781
Cl	GS-1S	1.3719
	GS-1S	1.3844
	GS-2P	1.1812
	GS-2P	2.5656
	GS-3D	1.8125
	GS-3D	2.3750
	GS-4F	2.5250
Ar	GS-1S	1.5688
	GS-1S	1.5734
	GS-2P	1.3609

Table C.7 – continued

System	Function	Exp.
	GS-2P	2.8422
	GS-3D	2.0656
	GS-3D	2.6148
	GS-4F	2.9375

Table C.8: Exponents for $3z$ ANO-GSn basis (continued on subsequent pages).

System	Function	Exp.
H	GS-1S	0.9250
	GS-3S	1.7437
	GS-2P	1.8594
	GS-3P	2.1437
	GS-3D	2.7094
He	GS-1S	2.6875
	GS-3S	4.4625
	GS-2P	2.5187
	GS-3P	3.0000
	GS-3D	3.8344
Li	GS-1S	0.5781
	GS-3S	0.7250
	GS-2P	1.6438
	GS-3P	0.8656
	GS-4P	1.4031
	GS-3D	0.7750
	GS-4D	1.2750

Table C.8 – continued

System	Function	Exp.
	GS-4F	1.1875
Be	GS-1S	0.6031
	GS-3S	1.0625
	GS-2P	2.0219
	GS-3P	0.9344
	GS-3D	1.0500
	GS-4D	1.7781
	GS-4F	1.5781
B	GS-1S	0.7875
	GS-3S	1.4094
	GS-2P	2.2500
	GS-3P	1.1156
	GS-3D	1.8094
	GS-4D	1.6344
	GS-4F	2.2172
C	GS-1S	0.9469
	GS-3S	1.7500
	GS-2P	2.9219
	GS-3P	1.3438
	GS-3D	1.9719
	GS-4D	2.1719
	GS-4F	2.7531
N	GS-1S	1.0312
	GS-3S	1.9000

Table C.8 – continued

System	Function	Exp.
	GS-2P	3.5063
	GS-3P	1.5688
	GS-3D	3.0094
	GS-4D	2.5469
	GS-4F	3.2594
O	GS-1S	1.0469
	GS-3S	2.4906
	GS-2P	3.8625
	GS-3P	1.5781
	GS-3D	2.5969
	GS-4D	2.9562
	GS-4F	3.5344
F	GS-1S	1.1375
	GS-3S	2.8906
	GS-2P	4.4125
	GS-3P	1.8469
	GS-3D	4.3156
	GS-4D	3.3312
	GS-4F	4.0594
Ne	GS-1S	1.4500
	GS-3S	3.3563
	GS-2P	5.0750
	GS-3P	2.1656
	GS-3D	5.0031

Table C.8 – continued

System	Function	Exp.
	GS-4D	3.8500
	GS-4F	4.8750
Na	GS-1S	0.5219
	GS-3S	0.6687
	GS-2P	0.5875
	GS-3P	0.7406
	GS-4P	1.2937
	GS-3D	0.6281
	GS-4D	1.0875
	GS-4F	1.0469
Mg	GS-1S	0.4125
	GS-3S	0.9125
	GS-2P	0.6781
	GS-3P	0.9656
	GS-3D	1.1125
	GS-4D	1.1531
	GS-4F	1.2437
Al	GS-1S	0.6281
	GS-3S	1.1156
	GS-2P	0.6188
	GS-3P	0.9250
	GS-3D	1.0969
	GS-4D	1.1719
	GS-4F	1.4688

Table C.8 – continued

System	Function	Exp.
Si	GS-1S	0.6188
	GS-3S	1.2531
	GS-2P	0.7531
	GS-3P	1.2031
	GS-3D	1.6187
	GS-4D	1.3594
	GS-4F	1.7969
P	GS-1S	0.7375
	GS-3S	1.4594
	GS-2P	0.8688
	GS-3P	1.4375
	GS-3D	1.9625
	GS-4D	1.6687
	GS-4F	2.0812
S	GS-1S	0.7875
	GS-3S	1.6687
	GS-2P	0.9156
	GS-3P	1.6313
	GS-3D	2.1969
	GS-4D	1.8469
	GS-4F	2.2844
Cl	GS-1S	0.8000
	GS-3S	1.8781
	GS-2P	1.0063

Table C.8 – continued

System	Function	Exp.
	GS-3P	1.7906
	GS-3D	2.3484
	GS-4D	2.1000
	GS-4F	2.5375
Ar	GS-1S	0.8094
	GS-3S	2.0875
	GS-2P	1.1156
	GS-3P	1.9500
	GS-3D	2.1281
	GS-4D	3.8687
	GS-4F	2.9312

C.3 Results

The total energies for the atoms hydrogen through argon using CCSD with the BFD, ANO-GS, and ANO-GSn bases are shown in Table C.9. The total energies for the homonuclear dimers of hydrogen through argon using CCSD with the BFD, ANO-GS, and ANO-GSn bases are shown in Table C.10. The total energies for Li, O, F, P, and S using several different electronic structure methods with the BFD, ANO-GS, and ANO-GSn bases are shown in Table C.11. The total energies for LiF, O_2, P_2, S_2, and SO_2 using several different electronic structure methods with the BFD, ANO-GS, and ANO-GSn bases are shown in Table C.12. The atomization energies for LiF, O_2, P_2, S_2, and SO_2 using several different

Table C.9: The total energies (in Hartrees) for the atoms hydrogen through argon using CCSD with the BFD, ANO-GS, and ANO-GSn bases. The results for the $2z$ ANO-GS/GSn are identical. In particular, there is no difference for lithium or sodium since their atoms are treated exactly in both cases.

	$2z$ BFD	$2z$ ANO-GS/GSn	$3z$ BFD	$3z$ ANO-GS	$3z$ ANO-GSn	$5z$ BFD
H	-0.499045	-0.500008	-0.499043	-0.500008	-0.500008	-0.499905
He	-2.878934	-2.897624	-2.898728	-2.901811	-2.901593	-2.902789
Li	-0.195611	-0.196326	-0.196093	-0.196326	-0.196326	-0.196315
Be	-1.000525	-1.008573	-1.008436	-1.009838	-1.009939	-1.009957
B	-2.608084	-2.606818	-2.615667	-2.616105	-2.616217	-2.617624
C	-5.408172	-5.406662	-5.425180	-5.426030	-5.426303	-5.429634
N	-9.759589	-9.761698	-9.788099	-9.789509	-9.790472	-9.796294
O	-15.828908	-15.833195	-15.879079	-15.881652	-15.884088	-15.896580
F	-24.091656	-24.099703	-24.159640	-24.164668	-24.166728	-24.186059
Ne	-34.899475	-34.911426	-34.972194	-34.990730	-34.992328	-35.018857
Na	-0.174227	-0.182144	-0.181799	-0.182144	-0.182144	-0.182034
Mg	-0.816857	-0.818098	-0.819095	-0.819352	-0.819531	-0.819679
Al	-1.928593	-1.928912	-1.935164	-1.935209	-1.935676	-1.936678
Si	-3.746217	-3.749091	-3.759962	-3.760189	-3.761051	-3.763306
P	-6.440629	-6.445748	-6.464683	-6.465607	-6.466748	-6.470915
S	-10.062475	-10.072356	-10.109733	-10.110522	-10.113578	-10.123942
Cl	-14.872848	-14.884950	-14.938980	-14.940167	-14.943319	-14.961472
Ar	-21.040743	-21.055925	-21.123540	-21.124919	-21.129124	-21.155927

electronic structure methods with the BFD, ANO-GS, and ANO-GSn bases are shown in Table C.13.

Table C.10: The total energies (in Hartrees) for the homonuclear dimers of hydrogen through argon using CCSD with the BFD, ANO-GS, and ANO-GSn bases. The results for the $2z$ ANO-GS/GSn are nearly identical. In particular, the lithium and sodium results differ by ~ 0.01 mH between the two basis sets.

	$2z$ BFD	$2z$ ANO-GS/GSn	$3z$ BFD	$3z$ ANO-GS	$3z$ ANO-GSn	$5z$ BFD
H_2	-1.170217	-1.170463	-1.173126	-1.173924	-1.173970	-1.174918
He_2	-5.757867	-5.795248	-5.797456	-5.803622	-5.803186	-5.805578
Li_2	-0.429057	-0.429487	-0.431324	-0.431277	-0.431339	-0.431475
Be_2	-1.994640	-2.012318	-2.013852	-2.017464	-2.017607	-2.018268
B_2	-5.294354	-5.300153	-5.318641	-5.320255	-5.320438	-5.325104
C_2	-10.993974	-11.005633	-11.043724	-11.047164	-11.047923	-11.059142
N_2	-19.821534	-19.837516	-19.905031	-19.911701	-19.912931	-19.934928
O_2	-31.803582	-31.818415	-31.922614	-31.930948	-31.934688	-31.966232
F_2	-48.209467	-48.230636	-48.361317	-48.370735	-48.376440	-48.419828
Ne_2	-69.798933	-69.822943	-69.944427	-69.981464	-69.984658	-70.037743
Na_2	-0.381541	-0.389296	-0.390139	-0.390741	-0.390815	-0.390813
Mg_2	-1.632893	-1.635623	-1.638339	-1.638985	-1.639314	-1.639841
Al_2	-3.889249	-3.891268	-3.908157	-3.908561	-3.909380	-3.912708
Si_2	-7.582780	-7.589569	-7.624903	-7.627154	-7.628772	-7.638618
P_2	-13.006848	-13.019340	-13.079636	-13.083205	-13.085725	-13.105334
S_2	-20.243451	-20.267665	-20.357425	-20.361329	-20.367137	-20.396481
Cl_2	-29.805005	-29.833665	-29.955047	-29.956971	-29.965052	-30.007852
Ar_2	-42.081394	-42.111987	-42.247056	-42.249899	-42.258331	-42.312045

Table C.11: Total energies (in Hartree) for Li, O, F, P, and S using several different electronic structure methods. Calculations are performed with the BFD, ANO-GS, and ANO-GSn bases. However, the $5z$ BFD* calculations do not include the G or H functions from the $5z$ BFD basis. All diffusion Monte Carlo (DMC) calculations are performed with a trial wavefunction obtained by optimizing Jastrow and orbital parameters via the linear method [19, 20, 21] in variational Monte Carlo. The DMC calculations are for a single-CSF reference (DMC-1CSF). For these systems, this is equivalent to a full-valence complete active space reference (DMC-FVCAS). A 0.01 H^{-1} time step is used for DMC calculations (continued on subsequent pages).

			RHF	B3LYP	CCSD	DMC-1CSF
Li	2z	BFD	-0.19561	-0.19754	-0.19561	-0.1963293(8)
Li	2z	ANO-GS	-0.19633	-0.19778	-0.19633	-0.1963299(5)
Li	2z	ANO-GSn	-0.19633	-0.19778	-0.19633	-0.1963304(8)
Li	3z	BFD	-0.19609	-0.19767	-0.19609	-0.1963303(4)
Li	3z	ANO-GS	-0.19633	-0.19781	-0.19633	-0.1963294(4)
Li	3z	ANO-GSn	-0.19633	-0.19781	-0.19633	-0.1963295(4)
Li	5z	BFD*	-0.19631	-0.19785	-0.19631	-0.1963306(3)
O	2z	BFD	-15.70594	-15.89680	-15.82891	-15.89316(9)
O	2z	ANO-GS/GSn	-15.70470	-15.89796	-15.83319	-15.89320(9)
O	3z	BFD	-15.70800	-15.89890	-15.87908	-15.89324(7)
O	3z	ANO-GS	-15.70785	-15.89924	-15.88165	-15.89312(7)
O	3z	ANO-GSn	-15.70799	-15.89925	-15.88409	-15.89321(7)
O	5z	BFD*	-15.70845	-15.89940	-15.89188	-15.89320(7)
F	2z	BFD	-23.93702	-24.19552	-24.09166	-24.18623(5)
F	2z	ANO-GS/GSn	-23.93594	-24.19760	-24.09970	-24.18621(5)
F	3z	BFD	-23.93822	-24.19875	-24.15964	-24.18642(7)
F	3z	ANO-GS	-23.93797	-24.19927	-24.16467	-24.18651(7)
F	3z	ANO-GSn	-23.93815	-24.19927	-24.16673	-24.18650(7)

Table C.11 – continued

			RHF	B3LYP	CCSD	DMC-1CSF
F	5z	BFD*	-23.93849	-24.19952	-24.17812	-24.18658(8)
P	2z	BFD	-6.35908	-6.44809	-6.44063	-6.47696(9)
P	2z	ANO-GS/GSn	-6.35899	-6.44816	-6.44575	-6.47698(9)
P	3z	BFD	-6.35908	-6.44841	-6.46468	-6.47705(5)
P	3z	ANO-GS	-6.35907	-6.44825	-6.46561	-6.47705(5)
P	3z	ANO-GSn	-6.35907	-6.44823	-6.46675	-6.47702(6)
P	5z	BFD*	-6.35908	-6.44850	-6.46905	-6.47697(5)
S	2z	BFD	-9.95531	-10.09844	-10.06248	-10.1318(1)
S	2z	ANO-GS/GSn	-9.95541	-10.09874	-10.07236	-10.1319(1)
S	3z	BFD	-9.95714	-10.09970	-10.10973	-10.13191(8)
S	3z	ANO-GS	-9.95661	-10.09981	-10.11052	-10.13191(7)
S	3z	ANO-GSn	-9.95727	-10.09986	-10.11358	-10.13191(8)
S	5z	BFD*	-9.95742	-10.10005	-10.11845	-10.13195(7)

Table C.12: Total energies (in Hartree) of several systems from G2 set [44] at their experimental geometries [35] using different electronic structure methods. Calculations are performed with the BFD, ANO-GS, and ANO-GSn bases. All diffusion Monte Carlo calculations are performed with a trial wavefunction obtained by optimizing Jastrow, orbital, and configuration state function (CSF) parameters (where applicable) via the linear method [19, 20, 21] in variational Monte Carlo. For each system, the DMC calculations are performed with both a single-CSF reference (DMC-1CSF) and full-valence complete active space reference (DMC-FVCAS). A 0.01 H^{-1} time step is used for DMC calculations.

			RHF	B3LYP	CCSD	DMC-1CSF	DMC FVCAS
LiF	2z	BFD	-24.27255	-24.59507	-24.47966	-24.6034(1)	-24.6073(1)
LiF	2z	ANO-GS	-24.27844	-24.60193	-24.49408	-24.60552(8)	-24.60829(7)
LiF	2z	ANO-GSn	-24.27946	-24.60327	-24.49568	-24.6057(2)	-24.60852(7)

Table C.12 – continued

			RHF	B3LYP	CCSD	DMC-1CSF	DMC FVCAS
LiF	3z	BFD	-24.28220	-24.60833	-24.56394	-24.60567(9)	-24.60985(6)
LiF	3z	ANO-GS	-24.28439	-24.61057	-24.57235	-24.6060(1)	-24.61005(7)
LiF	3z	ANO-GSn	-24.28430	-24.61052	-24.57310	-24.6059(1)	-24.6102(1)
LiF	5z	BFD*	-24.28550	-24.61160	-24.58740	-24.6060(1)	-24.61026(7)
O_2	2z	BFD	-31.43491	-31.96283	-31.80358	-31.9601(1)	-31.9715(1)
O_2	2z	ANO-GS/GSn	-31.43861	-31.97019	-31.81842	-31.9642(1)	-31.97439(9)
O_2	3z	BFD	-31.45438	-31.97639	-31.92261	-31.9697(1)	-31.97562(9)
O_2	3z	ANO-GS	-31.45616	-31.97892	-31.93095	-31.9709(1)	-31.97695(7)
O_2	3z	ANO-GSn	-31.45657	-31.97930	-31.93469	-31.9712(1)	-31.97731(8)
O_2	5z	BFD*	-31.45869	-31.98096	-31.95414	-31.9721(1)	-31.97816(5)
P_2	2z	BFD	-12.76276	-13.06483	-13.00685	-13.12686(9)	-13.13125(8)
P_2	2z	ANO-GS/GSn	-12.77082	-13.07018	-13.01934	-13.12796(9)	-13.13256(8)
P_2	3z	BFD	-12.77495	-13.07404	-13.07964	-13.1283(1)	-13.13272(8)
P_2	3z	ANO-GS	-12.77794	-13.07675	-13.08320	-13.13017(9)	-13.13472(8)
P_2	3z	ANO-GSn	-12.77785	-13.07642	-13.08573	-13.1299(1)	-13.13456(8)
P_2	5z	BFD*	-12.77890	-13.07769	-13.09576	-13.13082(9)	-13.13535(4)
S_2	2z	BFD	-19.97287	-20.34439	-20.24345	-20.4186(1)	-20.4193(1)
S_2	2z	ANO-GS/GSn	-19.98230	-20.35322	-20.26766	-20.4212(1)	-20.42199(9)
S_2	3z	BFD	-19.98990	-20.35785	-20.35743	-20.4218(1)	-20.4225(1)
S_2	3z	ANO-GS	-19.99188	-20.36111	-20.36133	-20.4237(1)	-20.4246(1)
S_2	3z	ANO-GSn	-19.99353	-20.36134	-20.36714	-20.4236(1)	-20.4246(1)
S_2	5z	BFD*	-19.99451	-20.36265	-20.38135	-20.4249(1)	-20.42569(9)
SO_2	2z	BFD	-41.48746	-42.22186	-42.00477	-42.2939(1)	-42.2983(1)
SO_2	2z	ANO-GS/GSn	-41.52259	-42.26068	-42.05531	-42.3126(1)	-42.3180(1)

Table C.12 – continued

			RHF	B3LYP	CCSD	DMC-1CSF	DMC FVCAS
SO_2	3z	BFD	-41.55800	-42.28125	-42.22455	-42.3171(1)	-42.3224(1)
SO_2	3z	ANO-GS	-41.56769	-42.29184	-42.24229	-42.32256(8)	-42.32781(6)
SO_2	3z	ANO-GSn	-41.56777	-42.29155	-42.24683	-42.32216(8)	-42.32734(6)
SO_2	5z	BFD*	-41.57326	-42.29628	-42.27913	-42.32436(8)	-42.33020(6)

Table C.13: Atomization energies (in kcal/mol) of several systems from G2 set [44] at their experimental geometries [35] using different electronic structure methods. Calculations are performed with the BFD, ANO-GS, and ANO-GSn bases. Calculated atomization energies are corrected with the experimental zero point energies [35, 45]. All diffusion Monte Carlo calculations are performed with a trial wavefunction obtained by optimizing Jastrow, orbital, and configuration state function (CSF) parameters (where applicable) via the linear method [19, 20, 21] in variational Monte Carlo. For each system, the DMC calculations are performed with both a single-CSF reference (DMC-1CSF) and full-valence complete active space reference (DMC-FVCAS). A 0.01 H^{-1} time step is used for DMC calculations.

			RHF	B3LYP	CCSD	DMC-1CSF	DMC FVCAS	Experiment
LiF	2z	BFD	86.5	125.5	119.4	137.28(7)	139.73(7)	138(2)
LiF	2z	ANO-GS	90.4	128.3	123.0	138.62(6)	140.36(5)	138(2)
LiF	2z	ANO-GSn	91.1	129.2	124.0	138.7(1)	140.51(5)	138(2)
LiF	3z	BFD	91.5	131.7	129.4	138.59(7)	141.21(6)	138(2)
LiF	3z	ANO-GS	92.9	132.7	131.3	138.74(8)	141.28(6)	138(2)
LiF	3z	ANO-GSn	92.7	132.6	130.5	138.68(8)	141.38(8)	138(2)
LiF	5z	BFD*	93.3	133.1	132.3	138.69(8)	141.37(7)	138(2)
O_2	2z	BFD	12.2	103.9	89.2	106.8(1)	114.0(1)	117.96(2)
O_2	2z	ANO-GS/GSn	16.1	107.1	93.1	109.3(1)	115.7(1)	117.96(2)
O_2	3z	BFD	21.8	109.8	100.9	112.7(1)	116.4(1)	117.96(2)
O_2	3z	ANO-GS	23.1	111.0	102.9	113.6(1)	117.4(1)	117.96(2)

Table C.13 – continued

			RHF	B3LYP	CCSD	DMC-1CSF	DMC FVCAS	Experiment
O_2	3z	ANO-GSn	23.2	111.2	102.2	113.7(1)	117.5(1)	117.96(2)
O_2	5z	BFD*	24.0	112.1	104.7	114.3(1)	118.08(9)	117.96(2)
P_2	2z	BFD	26.9	104.7	77.7	107.4(1)	110.2(1)	116.1(5)
P_2	2z	ANO-GS/GSn	32.0	108.0	79.1	108.1(1)	111.0(1)	116.1(5)
P_2	3z	BFD	34.5	110.1	93.2	108.20(9)	110.97(8)	116.1(5)
P_2	3z	ANO-GS	36.4	112.0	94.3	109.37(8)	112.23(8)	116.1(5)
P_2	3z	ANO-GSn	36.4	111.8	94.4	109.2(1)	112.16(9)	116.1(5)
P_2	5z	BFD*	37.0	112.3	97.8	109.88(8)	112.72(7)	116.1(5)
S_2	2z	BFD	38.0	91.5	73.3	96.2(1)	96.7(1)	100.66(7)
S_2	2z	ANO-GS/GSn	43.8	96.7	76.1	97.7(1)	98.2(1)	100.66(7)
S_2	3z	BFD	46.4	98.4	85.5	98.1(1)	98.5(1)	100.66(7)
S_2	3z	ANO-GS	48.3	100.3	87.0	99.3(1)	99.9(1)	100.66(7)
S_2	3z	ANO-GSn	48.5	100.4	86.8	99.2(1)	99.9(1)	100.66(7)
S_2	5z	BFD*	49.0	101.0	89.6	100.0(1)	100.5(1)	100.66(7)
SO_2	2z	BFD	71.1	202.6	174.2	231.5(1)	234.2(1)	254.0(2)
SO_2	2z	ANO-GS/GSn	94.7	225.3	194.3	243.1(1)	246.5(1)	254.0(2)
SO_2	3z	BFD	111.7	236.5	219.5	245.9(1)	249.2(1)	254.0(2)
SO_2	3z	ANO-GS	118.3	242.6	226.9	249.4(1)	252.7(1)	254.0(2)
SO_2	3z	ANO-GSn	117.7	242.4	224.8	249.1(1)	252.3(1)	254.0(2)
SO_2	5z	BFD*	120.5	245.1	232.2	250.4(1)	254.1(1)	254.0(2)

APPENDIX D

APPENDIX FOR CHAPTER 4

D.1 Hydrogen Basis Sets

The s contraction for the hydrogen basis sets is given in Table D.1. The Gauss-Slater (GS) primitives for the double-zeta ($2z$) atomic natural orbital GS (ANO-GS) basis for hydrogen are provided in Table D.2. The GS primitives for the triple-zeta ($3z$) ANO-GS basis for hydrogen are provided in Table D.3. The Gaussian primitives for the quintuple-zeta ($5z$) Gaussian basis for hydrogen are provided in Table D.4. Each basis was constructed using the methods of Ref. [25].

Table D.1: s contraction for hydrogen basis sets. The contraction was constructed using the method of Ref. [25].

Exponent	Coefficient
0.05559917	0.070858475
0.11675826	0.283134279
0.24519235	0.347654627
0.51490394	0.233232914
1.08129828	0.119346810
2.27072639	0.053366645
4.76852541	0.022054957
10.01390336	0.011743843
21.02919706	0.000820785

Table D.2: Primitives for $2z$ ANO-GS basis for hydrogen. The primitives were constructed using the method of Ref. [25].

Function	Exp.
GS-1S	1.946875
GS-2P	1.968750

Table D.3: Primitives for $3z$ ANO-GS basis for hydrogen. The primitives were constructed using the method of Ref. [25].

Function	Exp.
GS-1S	0.925000
GS-1S	1.968750
GS-2P	1.887500
GS-2P	2.553125
GS-3D	2.659375

Table D.4: Primitives for $5z$ Gaussian basis for hydrogen. The primitives were constructed using the method of Ref. [25].

Function	Exp.
1S	0.0692135
1S	0.1736131
1S	0.4543641
1S	1.3072524
2P	0.2370069
2P	0.6258464

Table D.4 – continued

Function	Exp.
2P	1.6114234
2P	5.1965679
3D	0.5106775
3D	1.3660613
3D	3.2652161
4F	0.6792030
4F	2.0042817

D.2 Setup and Reference Data

The geometries, zero point energies, and experimental atomization energies for the molecules of the G2 set [44] are provided in Table D.5.

Table D.5: Geometries, zero point energies, and experimental atomization energies for the molecules of the G2 set [44]. Some experimental error bars were not available (N/A).

System	Geo.	ZPE (kcal/mol)	Exp. Atomization En. (kcal/mol)
LiH	[51]	1.99 [51]	56 ± 0.01 [45]
BeH	[51]	2.92 [51]	46.9 ± 0.01 [45]
CH	[51]	4.04 [51]	79.9 ± 0.02 [45]
CH_2 ($^3B_1$)	[51]	10.55 [51]	$179.6 \pm$ N/A [45]
CH_2 ($^1A_1$)	[51]	10.29 [51]	$170.6 \pm$ N/A [45]
CH_3	[51]	18.55 [51]	289.3 ± 0.2 [45]
CH_4	[51]	27.74 [51]	392.5 ± 0.1 [45]

Table D.5 – continued

System	Geo.	ZPE (kcal/mol)	Expt. Atomization En. (kcal/mol)
NH	[51]	4.64 [51]	79 ± 0.4 [45]
NH_2	[51]	11.84 [51]	170 ± 0.3 [45]
NH_3	[51]	21.33 [51]	276.7 ± 0.1 [45]
OH	[51]	5.29 [51]	101.4 ± 0.3 [45]
H_2O	[51]	13.26 [51]	219.35 ± 0.01 [45]
HF	[51]	5.86 [51]	135.2 ± 0.2 [45]
SiH_2 ($^1A_1$)	[52]	7.3 [45]	144.4 ± 0.7 [45]
SiH_2 ($^3B_1$)	[53]	7.5 [45]	123.4 ± 0.7 [45]
SiH_3	[52]	13.2 [45]	213.8 ± 1.2 [45]
SiH_4	[52]	19.4 [45]	302.6 ± 0.5 [45]
PH_2	[52]	8.4 [45]	144.7 ± 0.6 [45]
PH_3	[51]	14.44 [51]	228.6 ± 0.4 [45]
H_2S	[51]	9.4 [51]	173.1 ± 0.2 [45]
HCl	[51]	4.24 [51]	102.24 ± 0.5 [45]
Li_2	[35]	0.5 [51]	23.9 ± 0.7 [45]
LiF	[35]	1.3 [45]	137.6 ± 2 [45]
C_2H_2	[51]	16.5 [51]	386.9 ± 0.2 [45]
C_2H_4	[51]	31.66 [51]	531.9 ± 0.1 [45]
C_2H_6	[51]	46.23 [51]	$666.3 \pm$ N/A [45]
CN	[51]	2.95 [51]	178.1 ± 2.4 [45]
HCN	[51]	9.95 [51]	301.7 ± 2 [45]
CO	[51]	3.09 [51]	256.2 ± 0.2 [45]
HCO	[51]	8.09 [51]	270.3 ± 2 [45]
H_2CO	[51]	16.52 [51]	357.2 ± 0.1 [45]

Table D.5 – continued

System	Geo.	ZPE (kcal/mol)	Expt. Atomization En. (kcal/mol)
H_3COH	[52]	31.72 [51]	480.8 ± N/A [45]
N_2	[51]	3.36 [51]	225.1 ± 0.4 [45]
N_2H_4	[52]	32.68 [51]	405.4 ± N/A [45]
NO	[51]	2.71 [51]	150.06 ± 0.04 [45]
O_2	[35]	2.25 [51]	117.96 ± 0.02 [45]
H_2O_2	[51]	16.44 [51]	252.3 ± N/A [45]
F_2	[35]	1.3 [51]	36.9 ± 0.1 [45]
CO_2	[51]	7.24 [51]	381.93 ± 0.01 [45]
Na_2	[52]	0.2 [45]	16.8 ± 0.3 [45]
Si_2	[35]	0.73 [51]	74 ± N/A [45]
P_2	[35]	1.11 [51]	116.1 ± 0.5 [45]
S_2	[35]	1.04 [51]	100.66 ± 0.07 [45]
Cl_2	[51]	0.8 [51]	57.18 ± 0.01 [45]
NaCl	[52]	0.5 [45]	97.3 ± 0.5 [45]
SiO	[51]	1.78 [51]	189.9 ± 2 [45]
CS	[35]	1.83 [51]	169.4 ± 6 [45]
SO	[51]	1.63 [51]	123.4 ± 0.3 [45]
ClO	[35]	1.22 [51]	63.42 ± 0.02 [45]
ClF	[51]	1.12 [51]	60.4 ± N/A [45]
Si_2H_6	[52]	30.5 [45]	500.1 ± N/A [45]
CH_3Cl	[51]	23.19 [51]	371 ± N/A [45]
H_3CSH	[52]	28.6 [45]	445.1 ± N/A [45]
HOCl	[51]	8.18 [51]	156.3 ± 0.5 [45]
SO_2	[35]	4.38 [51]	254 ± 0.2 [45]

D.3 Raw Data

D.3.1 Locality Approximation

This section contains the raw data for those systems handled with the locality approximation [55]. The diffusion Monte Carlo (DMC) total energies of the molecules from the G2 set and their atoms for a single determinant Slater-Jastrow (SJ) trial wavefunction composed of Hartree-Fock (HF) orbitals are given in Table D.6. The DMC total energies of the molecules from the G2 set and their atoms for a single determinant SJ trial wavefunction composed of variational Monte Carlo (VMC) optimized orbitals are given in Table D.7. The DMC total energies of the molecules from the G2 set and their atoms for a CAS SJ trial wavefunction with an s and p active space are given in Table D.8. The DMC total energies of the phosphorus containing molecules from the G2 set and their atoms for a single determinant SJ trial wavefunction composed of VMC optimized orbitals, a CAS SJ trial wavefunction with an s and p active space, and a CAS SJ trial wavefunction with an s, p, and d active space are given in Table D.9.

Table D.6: DMC total energies of the molecules from the G2 set and their atoms for a single determinant SJ trial wavefunction composed of HF orbitals. Energies are in Hartrees. Error bar, which is shown in parentheses, is for the last digit.

System	$2z$	$3z$	$5z$
H	-0.5000006(1)	-0.5000006(1)	-0.5000006(1)
Li	-0.1963297(5)	-0.1963297(7)	-0.196329(2)
C	-5.42251(7)	-5.42233(8)	-5.42236(7)
N	-9.79127(7)	-9.79109(5)	-9.79123(6)
O	-15.89289(8)	-15.89228(8)	-15.89254(8)

Table D.6 – continued

System	$2z$	$3z$	$5z$
F	-24.18633(8)	-24.18594(7)	-24.18609(7)
Na	-0.1821457(7)	-0.1821441(7)	-0.182145(2)
Si	-3.76661(5)	-3.76662(5)	-3.76654(5)
P	-6.47660(7)	-6.47648(5)	-6.47645(7)
S	-10.13125(9)	-10.13103(9)	-10.1311(1)
Cl	-14.97223(6)	-14.97168(6)	-14.97179(6)
CH	-6.05266(7)	-6.05399(9)	-6.05422(8)
CH_2 ($^1A_1$)	-6.7061(1)	-6.70751(9)	-6.70792(9)
CH_3	-7.41421(9)	-7.41476(8)	-7.41490(8)
CH_4	-8.09421(9)	-8.09441(9)	-8.09451(8)
NH	-10.41946(8)	-10.42028(7)	-10.42068(8)
NH_2	-11.07601(9)	-11.07774(9)	-11.07834(9)
NH_3	-11.7593(1)	-11.76181(8)	-11.76266(9)
OH	-16.56031(9)	-16.56126(9)	-16.56142(9)
H_2O	-17.25949(8)	-17.26123(8)	-17.26175(9)
HF	-24.90895(8)	-24.91024(7)	-24.91083(8)
SiH_2 ($^1A_1$)	-5.00856(7)	-5.00925(7)	-5.00952(7)
SiH_2 ($^3B_1$)	-4.97656(7)	-4.97723(7)	-4.97709(6)
SiH_3	-5.6280(1)	-5.62910(8)	-5.62906(9)
SiH_4	-6.28261(9)	-6.28418(8)	-6.28428(7)
PH_2	-7.71744(8)	-7.71798(8)	-7.71812(8)
PH_3	-8.35619(7)	-8.35676(7)	-8.35696(7)
H_2S	-11.41932(8)	-11.41962(7)	-11.42012(7)
HCl	-15.64183(8)	-15.64178(8)	-15.64211(8)

Table D.6 – continued

System	$2z$	$3z$	$5z$
Li_2	-0.43045(2)	-0.43053(3)	-0.43053(2)
C_2H_4	-13.74229(8)	-13.7433(1)	-13.74348(8)
C_2H_6	-14.98415(8)	-14.9851(1)	-14.9850(1)
CO	-21.7192(1)	-21.7232(1)	-21.7238(1)
H_2CO	-22.9038(1)	-22.9080(1)	-22.9082(1)
H_3COH	-24.12892(9)	-24.13207(8)	-24.13271(8)
N_2	-19.9296(1)	-19.9333(1)	-19.9338(1)
N_2H_4	-22.26561(9)	-22.27033(9)	-22.27206(8)
NO	-25.9062(1)	-25.9108(1)	-25.9117(1)
O_2	-31.9642(1)	-31.9679(1)	-31.9691(1)
H_2O_2	-33.1973(1)	-33.2013(1)	-33.2025(1)
F_2	-48.4134(1)	-48.4146(1)	-48.41757(9)
CO_2	-37.8168(1)	-37.8228(1)	-37.8235(1)
Na_2	-0.39102(2)	-0.39106(2)	-0.39104(2)
Si_2	-7.6498(1)	-7.6505(1)	-7.65031(9)
P_2	-13.1256(1)	-13.1273(1)	-13.1275(1)
S_2	-20.4191(1)	-20.4207(1)	-20.4214(1)
Cl_2	-30.0309(1)	-30.0305(1)	-30.03165(9)
SiO	-19.9535(1)	-19.96031(9)	-19.96104(8)
CS	-15.8151(1)	-15.8169(1)	-15.8185(1)
SO	-26.2105(1)	-26.2159(1)	-26.21669(9)
ClO	-30.9457(1)	-30.9475(1)	-30.9486(1)
ClF	-39.2434(1)	-39.24608(9)	-39.2478(1)
Si_2H_6	-11.3836(1)	-11.3857(1)	-11.38579(9)

Table D.6 – continued

System	$2z$	$3z$	$5z$
CH_3Cl	-22.5236(1)	-22.5245(1)	-22.5249(1)
H_3CSH	-18.30598(9)	-18.3071(1)	-18.30752(9)
HOCl	-31.61547(9)	-31.6179(1)	-31.6193(1)
SO_2	-42.2986(1)	-42.3125(1)	-42.3152(1)

Table D.7: The DMC total energies of the molecules from the G2 set and their atoms for a single determinant SJ trial wavefunction composed of VMC optimized orbitals. Energies are in Hartrees. Error bar, which is shown in parentheses, is for the last digit.

System	$2z$	$3z$	$5z$
H	-0.5000005(1)	-0.5000006(1)	-0.5000005(1)
Li	-0.1963295(4)	-0.1963292(4)	-0.1963296(3)
C	-5.42288(7)	-5.42292(7)	-5.42296(6)
N	-9.79213(6)	-9.79230(6)	-9.79229(6)
O	-15.8930(1)	-15.89305(8)	-15.89310(7)
F	-24.18619(5)	-24.18651(7)	-24.18636(8)
Na	-0.1821438(7)	-0.1821446(4)	-0.1821438(4)
Si	-3.76682(5)	-3.76676(4)	-3.76683(4)
P	-6.47688(9)	-6.47699(5)	-6.47686(5)
S	-10.1317(1)	-10.13184(8)	-10.13189(7)
Cl	-14.97253(6)	-14.97252(6)	-14.97261(6)
CH	-6.05388(8)	-6.05473(7)	-6.05476(9)
CH_2 ($^1A_1$)	-6.70711(9)	-6.70861(8)	-6.70860(7)
CH_3	-7.41489(7)	-7.41544(9)	-7.41549(7)
CH_4	-8.09523(7)	-8.09523(8)	-8.09523(8)

Table D.7 – continued

System	$2z$	$3z$	$5z$
NH	-10.42146(8)	-10.42200(8)	-10.42186(7)
NH_2	-11.07824(9)	-11.07960(8)	-11.07962(8)
NH_3	-11.76205(8)	-11.76368(7)	-11.76378(7)
OH	-16.56188(9)	-16.56243(8)	-16.56264(9)
H_2O	-17.26195(8)	-17.26298(8)	-17.26313(8)
HF	-24.91098(7)	-24.91170(6)	-24.91175(7)
SiH_2 ($^1A_1$)	-5.00947(5)	-5.01034(5)	-5.01055(4)
SiH_2 ($^3B_1$)	-4.97846(4)	-4.97932(3)	-4.97952(3)
SiH_3	-5.62975(6)	-5.63077(6)	-5.63097(5)
SiH_4	-6.28405(6)	-6.28526(5)	-6.28548(5)
PH_2	-7.71870(8)	-7.71952(6)	-7.71994(6)
PH_3	-8.35790(7)	-8.35892(5)	-8.35921(6)
H_2S	-11.42111(8)	-11.42203(7)	-11.42238(7)
HCl	-15.64294(7)	-15.64345(7)	-15.64374(7)
Li_2	-0.43069(2)	-0.43069(5)	-0.43077(4)
C_2H_4	-13.74413(7)	-13.74453(6)	-13.74477(7)
C_2H_6	-14.98622(7)	-14.98643(7)	-14.98640(6)
CO	-21.72262(9)	-21.72572(9)	-21.72617(9)
H_2CO	-22.9080(1)	-22.91048(9)	-22.91083(9)
H_3COH	-24.13189(8)	-24.13417(7)	-24.1348(1)
N_2	-19.93256(9)	-19.93600(9)	-19.93638(9)
N_2H_4	-22.26927(9)	-22.27373(7)	-22.27469(8)
NO	-25.9119(1)	-25.91612(9)	-25.91693(9)
O_2	-31.9682(1)	-31.9707(1)	-31.9719(1)

Table D.7 – continued

System	$2z$	$3z$	$5z$
H_2O_2	-33.2030(1)	-33.20627(8)	-33.20778(9)
F_2	-48.4216(1)	-48.42302(9)	-48.42522(8)
CO_2	-37.8259(1)	-37.8287(1)	-37.8291(1)
Na_2	-0.39104(1)	-0.39104(4)	-0.39106(2)
Si_2	-7.65138(6)	-7.65274(8)	-7.65279(7)
P_2	-13.12785(9)	-13.1300(1)	-13.13077(9)
S_2	-20.4211(1)	-20.4236(1)	-20.4248(1)
Cl_2	-30.03397(9)	-30.03514(9)	-30.03684(8)
SiO	-19.95896(9)	-19.96336(9)	-19.96385(7)
CS	-15.81932(9)	-15.82302(9)	-15.82387(8)
SO	-26.21600(9)	-26.22047(9)	-26.22172(9)
ClO	-30.9600(1)	-30.96312(9)	-30.96463(9)
ClF	-39.24935(9)	-39.25234(7)	-39.25406(8)
Si_2H_6	-11.38585(8)	-11.38818(7)	-11.38840(7)
CH_3Cl	-22.5259(1)	-22.5266(1)	-22.5275(1)
H_3CSH	-18.30852(9)	-18.30988(8)	-18.31105(7)
HOCl	-31.62113(9)	-31.62387(9)	-31.62501(8)
SO_2	-42.3125(1)	-42.32256(8)	-42.32431(8)

Table D.8: The DMC total energies of the molecules from the G2 set and their atoms for a CAS SJ trial wavefunction with an s and p active space. Energies are in Hartrees. Error bar, which is shown in parentheses, is for the last digit.

System	$3z$
H	-0.5000006(1)
Li	-0.1963292(4)

Table D.8 – continued

System	$3z$
C	-5.42823(6)
N	-9.79230(6)
O	-15.89305(8)
F	-24.18651(7)
Na	-0.1821446(4)
Si	-3.76711(4)
P	-6.47699(5)
S	-10.13184(8)
Cl	-14.97252(6)
CH	-6.06239(6)
CH_2 ($^1A_1$)	-6.71713(6)
CH_3	-7.41764(6)
CH_4	-8.09730(6)
NH	-10.42390(7)
NH_2	-11.08219(7)
NH_3	-11.76623(6)
OH	-16.56411(8)
H_2O	-17.26521(7)
HF	-24.91286(7)
SiH_2 ($^1A_1$)	-5.01238(4)
SiH_2 ($^3B_1$)	-4.97970(3)
SiH_3	-5.63114(4)
SiH_4	-6.28571(4)
PH_2	-7.72051(5)

Table D.8 – continued

System	$3z$
PH_3	-8.36003(5)
H_2S	-11.42305(7)
HCl	-15.64382(5)
Li_2	-0.431584(5)
C_2H_4	-13.75239(8)
C_2H_6	-14.99037(9)
CO	-21.73788(7)
H_2CO	-22.9190(1)
H_3COH	-24.14011(5)
N_2	-19.95032(7)
N_2H_4	-22.27934(5)
NO	-25.92787(8)
O_2	-31.97692(9)
H_2O_2	-33.21410(9)
F_2	-48.43320(8)
CO_2	-37.83890(7)
Na_2	-0.39106(2)
Si_2	-7.65485(7)
P_2	-13.13441(8)
S_2	-20.42452(9)
Cl_2	-30.03679(9)
SiO	-19.96885(7)
CS	-15.83223(6)
SO	-26.22235(9)

Table D.8 – continued

System	$3z$
ClO	-30.96605(8)
ClF	-39.25596(7)
Si_2H_6	-11.38838(7)
CH_3Cl	-22.52839(7)
H_3CSH	-18.3128(1)
HOCl	-31.6278(1)
SO_2	-42.33049(6)

Table D.9: The DMC total energies of the phosphorus containing molecules from the G2 set and their atoms for a single determinant SJ trial wavefunction composed of VMC optimized orbitals, a CAS SJ trial wavefunction with an s and p active space, and a CAS SJ trial wavefunction with an s, p, and d active space. Energies are in Hartrees. Error bar, which is shown in parentheses, is for the last digit.

System	1CSF	s, p CAS	s, p, d CAS
H	-0.5000006(1)	-0.5000006(1)	-0.5000006(1)
P	-6.47699(5)	-6.47699(5)	-6.47753(5)
PH_2	-7.71952(6)	-7.72051(5)	-7.72183(5)
PH_3	-8.35892(5)	-8.36003(5)	-8.36145(6)
P_2	-13.1300(1)	-13.13441(8)	-13.13814(7)

D.3.2 T-Moves

This section contains the raw data for those systems handled with the size-consistent version of T-Moves [56]. The DMC total energies of the molecules from the G2 set and their atoms for a single determinant SJ trial wavefunction composed of HF orbitals are given in Table D.10. The DMC total energies of the molecules from the G2 set and their atoms for a single determinant SJ trial wavefunction composed of VMC optimized orbitals are given in Table D.11. The DMC total energies of the molecules from the G2 set and their atoms for a CAS SJ trial wavefunction with an s and p active space are given in Table D.12.

Table D.10: DMC total energies of the molecules from the G2 set and their atoms for a single determinant SJ trial wavefunction composed of HF orbitals. Energies are in Hartrees. Error bar, which is shown in parentheses, is for the last digit.

System	$2z$	$3z$	$5z$
H	-0.5000006(1)	-0.5000006(1)	-0.5000006(1)
Li	-0.1963297(5)	-0.1963297(7)	-0.196329(2)
Be	-1.00918(5)	-1.00928(4)	-1.00926(5)
C	-5.42211(7)	-5.42206(7)	-5.42194(6)
N	-9.79121(5)	-9.79092(6)	-9.79123(6)
O	-15.89194(8)	-15.89176(8)	-15.89184(7)
F	-24.18514(7)	-24.18515(7)	-24.18527(7)
Na	-0.1821457(7)	-0.1821441(7)	-0.182145(2)
Cl	-14.97154(6)	-14.97117(6)	-14.97115(7)
LiH	-0.78787(3)	-0.78800(2)	-0.78799(1)
BeH	-1.58814(6)	-1.58861(4)	-1.58849(5)
CH_2 ($^3B_1$)	-6.72719(9)	-6.72762(7)	-6.72777(8)

Table D.10 – continued

System	$2z$	$3z$	$5z$
LiF	-24.60333(8)	-24.60401(8)	-24.60477(7)
C_2H_2	-12.4878(1)	-12.48906(9)	-12.48920(7)
CN	-15.48083(8)	-15.4833(1)	-15.48371(8)
HCN	-16.20259(8)	-16.20468(8)	-16.20494(9)
HCO	-22.24632(9)	-22.25065(8)	-22.25173(9)
NaCl	-15.31021(6)	-15.31104(8)	-15.31135(7)

Table D.11: The DMC total energies of the molecules from the G2 set and their atoms for a single determinant SJ trial wavefunction composed of VMC optimized orbitals. Energies are in Hartrees. Error bar, which is shown in parentheses, is for the last digit.

System	$2z$	$3z$	$5z$
H	-0.5000005(1)	-0.5000006(1)	-0.5000005(1)
Li	-0.1963295(4)	-0.1963292(4)	-0.1963296(3)
Be	-1.00921(5)	-1.00916(5)	-1.00924(4)
C	-5.42244(6)	-5.42258(5)	-5.42245(5)
N	-9.79214(6)	-9.79210(6)	-9.79218(6)
O	-15.89241(9)	-15.89249(7)	-15.89272(7)
F	-24.18526(5)	-24.18535(7)	-24.18548(7)
Na	-0.1821438(7)	-0.1821446(4)	-0.1821438(4)
Cl	-14.97209(6)	-14.97204(6)	-14.97212(5)
LiH	-0.78804(1)	-0.78804(1)	-0.788051(9)
BeH	-1.58833(4)	-1.58887(3)	-1.58883(3)
CH_2 ($^3B_1$)	-6.72867(7)	-6.72921(7)	-6.72940(6)
LiF	-24.60451(8)	-24.60473(8)	-24.60479(8)

Table D.11 – continued

System	$2z$	$3z$	$5z$
C_2H_2	-12.4900(1)	-12.49043(7)	-12.49048(8)
CN	-15.49091(9)	-15.49296(8)	-15.49339(8)
HCN	-16.20593(8)	-16.20711(7)	-16.20741(7)
HCO	-22.25183(8)	-22.25561(6)	-22.25607(7)
NaCl	-15.31167(5)	-15.31196(6)	-15.31202(7)

Table D.12: The DMC total energies of the molecules from the G2 set and their atoms for a CAS SJ trial wavefunction with an s and p active space. Energies are in Hartrees. Error bar, which is shown in parentheses, is for the last digit.

System	$3z$
H	-0.5000006(1)
Li	-0.1963292(4)
Be	-1.010186(5)
C	-5.42775(5)
N	-9.79230(6)
O	-15.89305(8)
F	-24.18651(7)
Na	-0.1821446(4)
Cl	-14.97252(6)
LiH	-0.788238(6)
BeH	-1.59014(2)
CH_2 ($^3B_1$)	-6.73142(8)
LiF	-24.61011(7)
C_2H_2	-12.50101(5)
CN	-15.50866(7)

Table D.12 – continued

System	$3z$
HCN	-16.22003(6)
HCO	-22.26378(7)
NaCl	-15.31334(6)

D.4 Variational Monte Carlo

This section contains the VMC data for the G2 set. The VMC total energies of the molecules from the G2 set and their atoms for a single determinant SJ trial wavefunction composed of HF orbitals are given in Table D.13. The VMC total energies of the molecules from the G2 set and their atoms for a single determinant SJ trial wavefunction composed of VMC optimized orbitals are given in Table D.14. The VMC total energies of the molecules from the G2 set and their atoms for a CAS SJ trial wavefunction with an s and p active space are given in Table D.15.

Table D.13: VMC total energies of the molecules from the G2 set and their atoms for a single determinant SJ trial wavefunction composed of HF orbitals. Energies are in Hartrees. Error bar, which is shown in parentheses, is for the last digit.

System	$2z$	$3z$	$5z$
H	-0.5000002(3)	-0.5000002(3)	-0.5000003(3)
Li	-0.1963192(5)	-0.1963192(5)	-0.196312(3)
Be	-1.00835(8)	-1.00843(8)	-1.0082(1)
C	-5.41531(5)	-5.41496(5)	-5.41513(2)
N	-9.78553(5)	-9.78506(5)	-9.78529(5)

Table D.13 – continued

System	$2z$	$3z$	$5z$
O	-15.88268(8)	-15.88277(8)	-15.88312(8)
F	-24.17448(9)	-24.17461(9)	-24.17472(9)
Na	-0.1821236(9)	-0.1821239(9)	-0.181993(6)
Si	-3.76315(2)	-3.76295(2)	-3.76298(2)
P	-6.47276(4)	-6.47260(4)	-6.47259(4)
S	-10.12444(5)	-10.12390(6)	-10.12434(6)
Cl	-14.96368(5)	-14.96340(5)	-14.96353(5)
LiH	-0.78727(5)	-0.78777(4)	-0.78784(3)
BeH	-1.58359(3)	-1.58575(2)	-1.58566(2)
CH	-6.04119(8)	-6.04419(7)	-6.04483(7)
CH_2 ($^3B_1$)	-6.71774(7)	-6.71864(7)	-6.71906(7)
CH_2 ($^1A_1$)	-6.69381(8)	-6.69681(7)	-6.69787(7)
CH_3	-7.40344(8)	-7.40419(8)	-7.40457(7)
CH_4	-8.08247(9)	-8.08310(8)	-8.08338(8)
NH	-10.40752(9)	-10.40999(9)	-10.41050(9)
NH_2	-11.0613(1)	-11.0651(1)	-11.06646(9)
NH_3	-11.74334(9)	-11.74816(8)	-11.74998(8)
OH	-16.5468(1)	-16.5490(1)	-16.5495(1)
H_2O	-17.24389(8)	-17.24777(7)	-17.24877(7)
HF	-24.89475(8)	-24.89747(7)	-24.89835(7)
SiH_2 ($^1A_1$)	-4.99763(9)	-4.99915(8)	-5.00002(8)
SiH_2 ($^3B_1$)	-4.96462(9)	-4.96581(8)	-4.96669(8)
SiH_3	-5.6151(1)	-5.61682(9)	-5.61788(9)
SiH_4	-6.27140(7)	-6.27356(7)	-6.27459(6)

Table D.13 – continued

System	$2z$	$3z$	$5z$
PH_2	-7.70464(7)	-7.70603(7)	-7.70671(6)
PH_3	-8.34136(9)	-8.34312(8)	-8.34374(8)
H_2S	-11.40510(6)	-11.40659(6)	-11.40795(6)
HCl	-15.63025(8)	-15.63067(8)	-15.63146(8)
Li_2	-0.42921(6)	-0.42957(8)	-0.42973(8)
LiF	-24.59015(9)	-24.5933(1)	-24.5941(1)
C_2H_2	-12.46639(9)	-12.46942(8)	-12.46978(8)
C_2H_4	-13.71321(7)	-13.71452(9)	-13.71543(7)
C_2H_6	-14.95497(8)	-14.95623(7)	-14.95686(7)
CN	-15.4528(1)	-15.4574(1)	-15.4578(1)
HCN	-16.1773(1)	-16.1815(1)	-16.1819(1)
CO	-21.6859(1)	-21.6938(1)	-21.6948(1)
HCO	-22.2121(1)	-22.2195(1)	-22.2209(1)
H_2CO	-22.86719(7)	-22.87324(7)	-22.87426(7)
H_3COH	-24.0890(1)	-24.0942(1)	-24.0957(1)
N_2	-19.8952(1)	-19.9019(1)	-19.9026(1)
N_2H_4	-22.22205(8)	-22.23083(8)	-22.23405(7)
NO	-25.8669(1)	-25.8745(1)	-25.8759(1)
O_2	-31.92106(9)	-31.9270(1)	-31.9288(1)
H_2O_2	-33.14808(9)	-33.15601(9)	-33.15818(9)
F_2	-48.3613(1)	-48.3642(1)	-48.3685(1)
CO_2	-37.7654(1)	-37.7788(1)	-37.7800(1)
Na_2	-0.39054(5)	-0.39066(4)	-0.39062(4)
Si_2	-7.63400(8)	-7.63601(7)	-7.63606(8)

Table D.13 – continued

System	$2z$	$3z$	$5z$
P_2	-13.10120(9)	-13.10558(9)	-13.10629(9)
S_2	-20.3917(1)	-20.3959(1)	-20.3968(1)
Cl_2	-29.99928(8)	-29.99951(8)	-30.00062(8)
NaCl	-15.30170(8)	-15.30289(8)	-15.30341(8)
SiO	-19.9247(1)	-19.9345(1)	-19.93647(9)
CS	-15.7859(1)	-15.79042(9)	-15.79377(9)
SO	-26.1736(1)	-26.1814(1)	-26.18343(9)
ClO	-30.9076(1)	-30.9113(1)	-30.9130(1)
ClF	-39.2039(1)	-39.2073(1)	-39.2092(1)
Si_2H_6	-11.35963(8)	-11.36288(7)	-11.36485(7)
CH_3Cl	-22.4938(1)	-22.4950(1)	-22.4961(1)
H_3CSH	-18.27368(9)	-18.27634(5)	-18.27741(9)
HOCl	-31.57616(8)	-31.58107(7)	-31.58258(7)
SO_2	-42.23410(8)	-42.2544(1)	-42.2591(1)

Table D.14: The VMC total energies of the molecules from the G2 set and their atoms for a single determinant SJ trial wavefunction composed of VMC optimized orbitals. Energies are in Hartrees. Error bar, which is shown in parentheses, is for the last digit.

System	$2z$	$3z$	$5z$
H	-0.5000006(3)	-0.5000004(3)	-0.5000005(3)
Li	-0.196321(1)	-0.196324(1)	-0.196325(3)
Be	-1.0084(1)	-1.00850(7)	-1.00837(8)
C	-5.41616(5)	-5.41625(5)	-5.41625(5)
N	-9.78640(6)	-9.78660(5)	-9.78647(5)

Table D.14 – continued

System	$2z$	$3z$	$5z$
O	-15.8832(1)	-15.88375(7)	-15.88404(8)
F	-24.17474(9)	-24.17534(9)	-24.17541(8)
Na	-0.1821236(9)	-0.182139(2)	-0.182141(3)
Si	-3.76433(2)	-3.76455(2)	-3.76457(2)
P	-6.4740(1)	-6.47435(3)	-6.47433(3)
S	-10.12575(4)	-10.12612(5)	-10.12647(5)
Cl	-14.96474(5)	-14.96515(5)	-14.96539(5)
LiH	-0.78780(2)	-0.78791(2)	-0.78795(2)
BeH	-1.58597(3)	-1.58761(2)	-1.58748(2)
CH	-6.04447(7)	-6.04668(7)	-6.04693(6)
CH_2 ($^3B_1$)	-6.72118(7)	-6.72238(6)	-6.72264(6)
CH_2 ($^1A_1$)	-6.69673(8)	-6.70002(7)	-6.70032(7)
CH_3	-7.40543(8)	-7.40662(7)	-7.40686(7)
CH_4	-8.08593(8)	-8.08620(8)	-8.08606(8)
NH	-10.41134(9)	-10.41286(8)	-10.41307(8)
NH_2	-11.0659(1)	-11.06868(9)	-11.06898(9)
NH_3	-11.74792(8)	-11.75212(8)	-11.75240(7)
OH	-16.5496(1)	-16.5510(1)	-16.5511(1)
H_2O	-17.24801(8)	-17.25070(7)	-17.25124(7)
HF	-24.89807(8)	-24.89954(7)	-24.89999(7)
SiH_2 ($^1A_1$)	-5.00102(8)	-5.00415(7)	-5.00467(6)
SiH_2 ($^3B_1$)	-4.97215(7)	-4.97453(6)	-4.97506(6)
SiH_3	-5.62188(9)	-5.62493(7)	-5.62549(7)
SiH_4	-6.27681(6)	-6.27942(5)	-6.27988(5)

Table D.14 – continued

System	$2z$	$3z$	$5z$
PH_2	-7.70860(7)	-7.71172(6)	-7.71254(5)
PH_3	-8.34631(8)	-8.34997(7)	-8.35084(7)
H_2S	-11.40914(6)	-11.41219(5)	-11.41365(5)
HCl	-15.63249(8)	-15.63411(8)	-15.63494(7)
Li_2	-0.43001(5)	-0.4300(1)	-0.43013(6)
LiF	-24.59338(9)	-24.5945(1)	-24.5946(1)
C_2H_2	-12.47251(8)	-12.47386(8)	-12.47424(8)
C_2H_4	-13.72072(7)	-13.72361(6)	-13.72434(6)
C_2H_6	-14.96347(7)	-14.96475(6)	-14.96536(6)
CN	-15.4670(1)	-15.4713(1)	-15.47249(9)
HCN	-16.1850(1)	-16.18778(9)	-16.18830(9)
CO	-21.6953(1)	-21.7026(1)	-21.7035(1)
HCO	-22.2243(1)	-22.23171(9)	-22.23330(9)
H_2CO	-22.87806(7)	-22.88477(6)	-22.88603(6)
H_3COH	-24.09967(9)	-24.10596(9)	-24.10826(9)
N_2	-19.9025(1)	-19.9103(1)	-19.91182(9)
N_2H_4	-22.23354(7)	-22.24360(7)	-22.24677(7)
NO	-25.8773(1)	-25.8862(1)	-25.8888(1)
O_2	-31.92806(7)	-31.9350(1)	-31.9392(1)
H_2O_2	-33.16077(9)	-33.16910(8)	-33.17382(8)
F_2	-48.3752(1)	-48.3822(1)	-48.3866(1)
CO_2	-37.7880(1)	-37.7942(1)	-37.7949(1)
Na_2	-0.39087(3)	-0.39082(4)	-0.39101(3)
Si_2	-7.63878(5)	-7.64202(7)	-7.64348(7)

Table D.14 – continued

System	$2z$	$3z$	$5z$
P_2	-13.10695(9)	-13.11214(8)	-13.11542(8)
S_2	-20.3953(1)	-20.4019(1)	-20.4059(1)
Cl_2	-30.0053(1)	-30.0087(1)	-30.01375(9)
NaCl	-15.30374(8)	-15.30468(7)	-15.30503(7)
SiO	-19.9338(1)	-19.94292(9)	-19.94447(9)
CS	-15.79466(9)	-15.80294(9)	-15.80576(8)
SO	-26.1840(1)	-26.19383(9)	-26.19689(9)
ClO	-30.9264(1)	-30.9332(1)	-30.9377(1)
ClF	-39.2141(1)	-39.22140(9)	-39.22499(9)
Si_2H_6	-11.36964(6)	-11.37477(6)	-11.37609(5)
CH_3Cl	-22.50049(9)	-22.50391(9)	-22.50675(9)
H_3CSH	-18.28229(9)	-18.28658(8)	-18.28989(8)
HOCl	-31.58679(7)	-31.59298(7)	-31.59752(7)
SO_2	-42.2567(1)	-42.2762(1)	-42.2831(1)

Table D.15: The VMC total energies of the molecules from the G2 set and their atoms for a CAS SJ trial wavefunction with an s and p active space. Energies are in Hartrees. Error bar, which is shown in parentheses, is for the last digit.

System	$3z$
H	-0.5000004(3)
Li	-0.196324(1)
Be	-1.010191(8)
C	-5.42342(4)
N	-9.78660(5)
O	-15.88375(7)

Table D.15 – continued

System	$3z$
F	-24.17534(9)
Na	-0.182139(2)
Si	-3.76508(2)
P	-6.47435(3)
S	-10.12612(5)
Cl	-14.96515(5)
LiH	-0.78818(1)
BeH	-1.58944(1)
CH	-6.05669(6)
CH_2 ($^3B_1$)	-6.72640(6)
CH_2 ($^1A_1$)	-6.71131(6)
CH_3	-7.41137(6)
CH_4	-8.09104(7)
NH	-10.41641(8)
NH_2	-11.07366(8)
NH_3	-11.75758(7)
OH	-16.5540(1)
H_2O	-17.25510(7)
HF	-24.90171(7)
SiH_2 ($^1A_1$)	-5.00884(4)
SiH_2 ($^3B_1$)	-4.97634(4)
SiH_3	-5.62691(5)
SiH_4	-6.28131(5)
PH_2	-7.71436(5)

Table D.15 – continued

System	$3z$
PH_3	-8.35357(6)
H_2S	-11.41515(5)
HCl	-15.63556(8)
Li_2	-0.43156(1)
LiF	-24.6016(1)
C_2H_2	-12.49172(6)
C_2H_4	-13.74004(5)
C_2H_6	-14.97618(8)
CN	-15.49602(8)
HCN	-16.20797(8)
CO	-21.72140(9)
HCO	-22.24735(8)
H_2CO	-22.90135(6)
H_3COH	-24.11924(8)
N_2	-19.93310(8)
N_2H_4	-22.25696(6)
NO	-25.9055(1)
O_2	-31.9489(1)
H_2O_2	-33.18541(3)
F_2	-48.4023(1)
CO_2	-37.8128(1)
Na_2	-0.39099(1)
Si_2	-7.64737(6)
P_2	-13.12193(8)

Table D.15 – continued

System	$3z$
S_2	-20.4049(1)
Cl_2	-30.0128(1)
NaCl	-15.30706(7)
SiO	-19.95375(9)
CS	-15.81754(3)
SO	-26.19852(9)
ClO	-30.93970(9)
ClF	-39.22866(9)
Si_2H_6	-11.37852(6)
CH_3Cl	-22.50929(9)
H_3CSH	-18.29484(7)
HOCl	-31.60189(7)
SO_2	-42.29272(9)

The deviation of the VMC atomization energies from experiment for a single determinant SJ trial wavefunction composed of HF orbitals is shown in Figure D.1. The MAD from experiment for the $2z$, $3z$, and $5z$ bases are 12.0 kcal/mol, 9.4 kcal/mol, and 8.9 kcal/mol, respectively. The deviation of the VMC atomization energies from experiment for a single determinant SJ trial wavefunction composed of VMC optimized orbitals is shown in Figure D.2. The MAD from experiment for the $2z$, $3z$, and $5z$ bases are 8.4 kcal/mol, 6.1 kcal/mol, and 5.2 kcal/mol, respectively. The deviation of the VMC atomization energies from experiment for the s and p valence CAS SJ trial wavefunctions is shown in Figure D.3. The $5z$ single determinant results are included to demonstrate the benefit

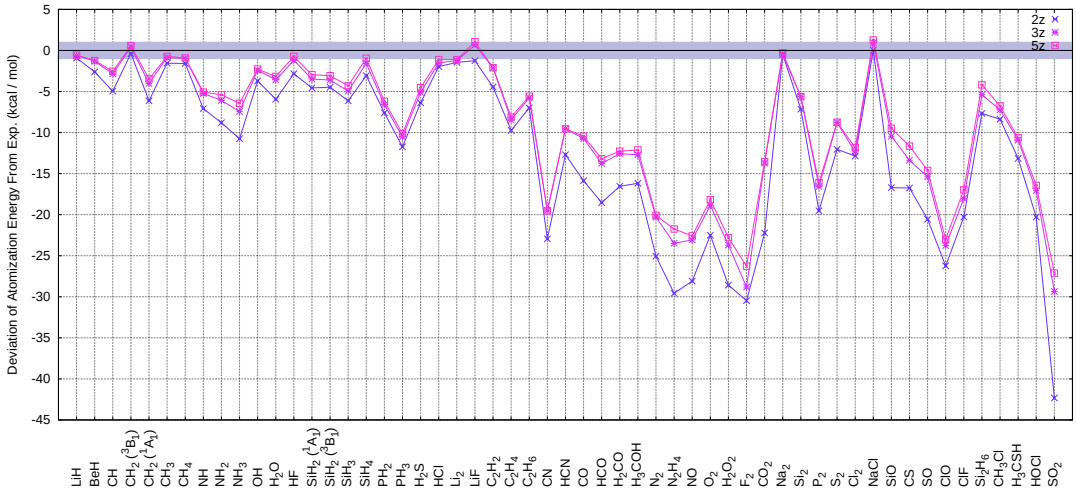

Figure D.1: Deviation of the VMC atomization energies from experiment for a single determinant SJ trial wavefunction composed of HF orbitals. The MAD from experiment for the $2z$, $3z$, and $5z$ bases are 12.0 kcal/mol, 9.4 kcal/mol, and 8.9 kcal/mol, respectively.

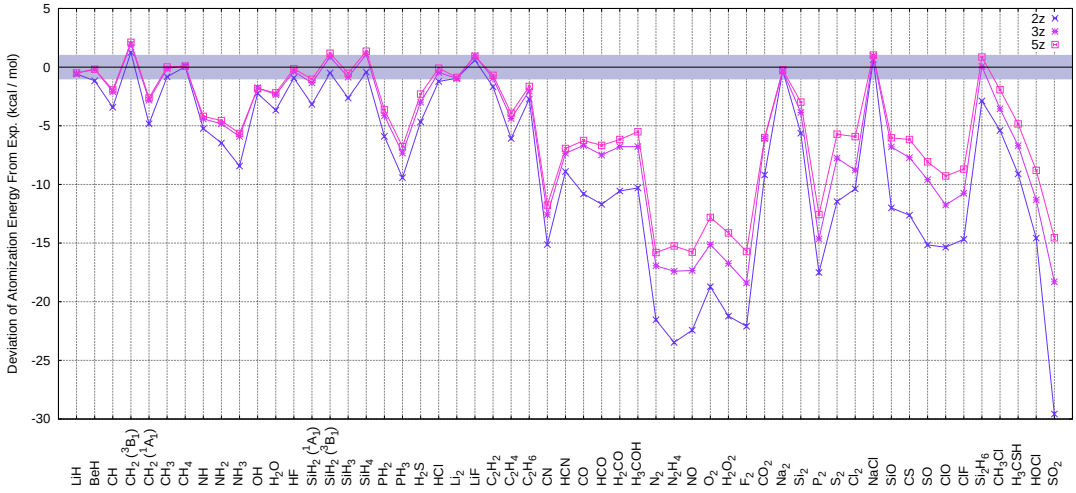

Figure D.2: Deviation of the VMC atomization energies from experiment for a single determinant SJ trial wavefunction composed of VMC optimized orbitals. The MAD from experiment for the $2z$, $3z$, and $5z$ bases are 8.4 kcal/mol, 6.1 kcal/mol, and 5.2 kcal/mol, respectively.

of using a CAS SJ trial wavefunction. This modest basis and CSF expansion results in a MAD from experiment of 2.9 kcal/mol.

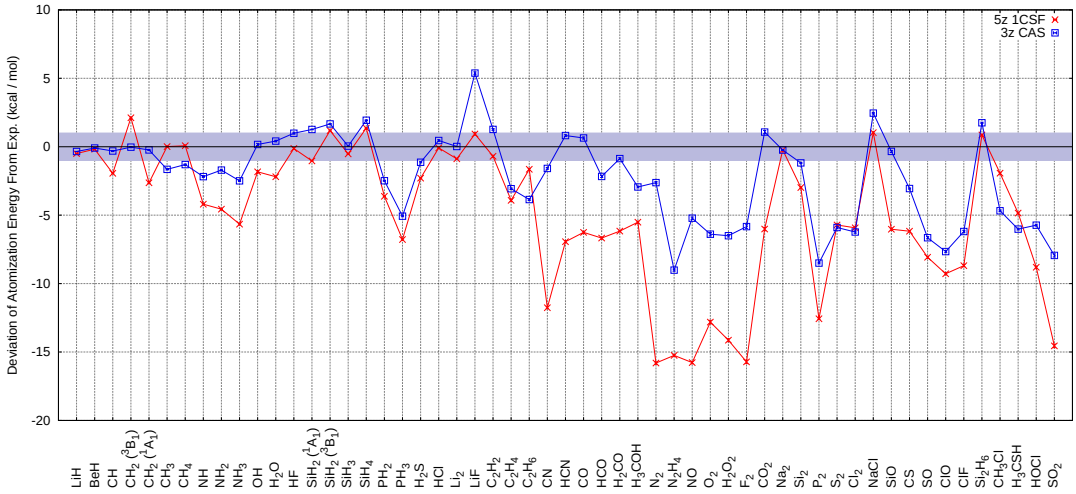

Figure D.3: Deviation of the VMC atomization energies from experiment for a single determinant SJ trial wavefunction composed of VMC optimized orbitals and a CAS SJ trial wavefunction. The MAD from experiment for the single determinant SJ trial wavefunction is 5.2 kcal/mol. The MAD from experiment for the CAS SJ trial wavefunction is 2.9 kcal/mol.

BIBLIOGRAPHY

[1] *Quantum Monte Carlo Methods in Physics and Chemistry*, edited by M. P. Nightingale and C. J. Umrigar, NATO ASI Ser. C 525 (Kluwer, Dordrecht, 1999)

[2] R. G. Parr and W. Yang, *Density-Functional Theory of Atoms and Molecules* (Oxford University Press, New York, 1989)

[3] T. Helgaker, P. Jorgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (John Wiley & Sons LTD, Chichester, England, 2000)

[4] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 3383 (2001)

[5] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. **21**, 1087 (1953)

[6] W. K. Hastings, Biometrika **57**, 97 (1970)

[7] F. R. Petruzielo, J. Toulouse, and C. J. Umrigar, J. Chem. Phys. **132**, 094109 (2010)

[8] P. C. Hariharan and J. A. Pople, Theor. Chem. Acta. **28**, 213 (1973)

[9] W. Hehre, R. Stewart, and J. Pople, J. Chem. Phys. **51**, 2657 (1969)

[10] T. Dunning Jr, J. Chem. Phys. **90**, 1007 (1989)

[11] R. Kendall, T. Dunning Jr, and R. Harrison, J. Chem. Phys. **96**, 6796 (1992)

[12] T. Kato, Comm. Pure Appl. Math **10**, 151177 (1957)

[13] S. Boys, Proc. R. Soc. A **200**, 542554 (1950)

[14] C. Umrigar, K. Wilson, and J. Wilkins, Phys. Rev. Lett. **60**, 17191722 (1988)

[15] C. Filippi and C. J. Umrigar, J. Chem. Phys. **105**, 213 (1996)

[16] C. J. Umrigar, M. P. Nightingale, and K. J. Runge, J. Chem. Phys. **99**, 2865 (1993)

[17] CHAMP, a quantum Monte Carlo program written by C. J. Umrigar, C. Filippi and Julien Toulouse. URL - http://www.physics.cornell.edu/ cyrus/champ.html

[18] M. Burkatzki, C. Filippi, and M. Dolg, J. Chem. Phys. **126**, 234105 (2007)

[19] J. Toulouse and C. J. Umrigar, J. Chem. Phys. **126**, 084102 (2007)

[20] J. Toulouse and C. J. Umrigar, J. Chem. Phys. **128**, 174101 (2008)

[21] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, Phys. Rev. Lett. **98**, 110201 (2007)

[22] K. L. Schuchardt, B. T. Didier, T. Elsethagen, L. Sun, V. Gurumoorthi, J. Chase, J. Li, and T. L. Windus, J. Chem. Inf. Model. **47**, 1045 (2007)

[23] D. Feller, J. Comp. Chem. **17**, 1571 (1996)

[24] M. W. Schmidt, J. A. Boatz, K. K. Baldridge, S. T. Elbert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, J. Comp. Chem. **14**, 1347 (1993)

[25] F. R. Petruzielo, J. Toulouse, and C. J. Umrigar, J. Chem. Phys. **134**, 064104 (2011)

[26] J. Almlöf and P. Taylor, J. Chem. Phys. **86**, 4070 (1987)

[27] P.-O. Widmark, P.-A. Malmqvist, and B. Roos, Theor. Chim. Acta. **77**, 291 (1990)

[28] P.-O. Widmark, B. Joakim, and B. Roos, Theor. Chim. Acta. **79**, 419 (1991)

[29] V. Veryazov, P.-O. Widmark, and B. O. Roos, Theor. Chim. Acta. **111**, 345 (2004)

[30] A. Becke, J. Chem. Phys. **98**, 5648 (1993)

[31] See Supplementary Material at http://dx.doi.org/10.1063/1.3551512 for atom specific basis sets, Gaussian fits of Gauss-Slater functions, and data used for making figures.

[32] K. P. Huber and G. Herzberg, *Constants of Diatomic Molecules, Molecular Spectra and Molecular Structure Vol 4* (Van Nostrand Reinhold Company, 1979)

[33] L. Gurvich, I. V. Veyts, and C. B. Alcock, *Thermodynamic Properties of Individual Substances, Fouth Edition* (Hemisphere Pub. Co., 1989)

[34] *NIST Chemistry WebBook, NIST Standard Reference Database Number 69*, edited by P. J. Linstrom and W. G. Mallard (NIST, Gaithersburg, MD, 2005)

[35] *NIST Computational Chemistry Comparison and Benchmark Database, NIST Standard Reference Database Number 101*, edited by R. D. Johnson (NIST, Gaithersburg, MD, 2010)

[36] R. Grisenti, W. Schollkopf, J. Toennies, G. Hegerfeldt, T. Kohler, and M. Stoll, Phys. Rev. Lett. **85**, 2284 (2000)

[37] V. E. Bondybey and J. H. English, J. Chem. Phys. **80**, 568 (1984)

[38] R. Aziz and M. Slaman, Chem. Phys. **130**, 187 (1989)

[39] Z. Fu, G. W. Lemire, G. A. Bishea, and M. D. Morse, J. Chem. Phys. **93**, 8420 (1990)

[40] P. R. Herman, P. E. Larocque, and B. P. Stoicheff, J. Chem. Phys. **89**, 4535 (1988)

[41] URL - http://www.burkatzki.com/pseudos/index.2.html

[42] K. Irikura, J. Phys. Chem. Ref. Data **36**, 389 (2007)

[43] Y. R. Luo, *Comprehensive Handbook of Chemical Bond Energies* (CRC Press, 2007)

[44] L. A. Curtiss, K. Raghavachari, G. W. Trucks, and J. A. Pople, J. Chem. Phys. **94**, 7221 (1991)

[45] D. Feller and K. A. Peterson, J. Chem. Phys. **110**, 8384 (1999)

[46] R. J. Needs, M. D. Towler, N. D. Drummond, and P. López Ríos, J. Phys. Condens. Matter **22**, 023201 (2010)

[47] J. C. Grossman, J. Chem. Phys. **117**, 1434 (2002)

[48] N. Nemec, M. D. Towler, and R. J. Needs, J. Chem. Phys. **132**, 034111 (2010)

[49] F. R. Petruzielo, J. Toulouse, and C. J. Umrigar, J. Chem. Phys. **136**, 124116 (2012)

[50] See supplementary material at http://dx.doi.org/10.1063/1.3697846 for hydrogen basis sets. The geometries, zero point energies, and experimental atomization energies for each molecule are included. Raw data for all calculations are also provided.

[51] D. Feller, K. A. Peterson, and D. A. Dixon, J. Chem. Phys. **129**, 204105 (2008)

[52] D. P. O'Neill and P. M. W. Gill, Mol. Phys. **103**, 763 (2005)

[53] A. Kalemos, T. H. Dunning Jr, and A. Mavridis, Mol. Phys. **102**, 2597 (2004)

[54] M. Nightingale and H. Blöte, Phys. Rev. B **33**, 659 (1986)

[55] L. Mitas, E. L. Shirley, and D. M. Ceperley, J. Chem. Phys. **95**, 3467 (1991)

[56] M. Casula, S. Moroni, S. Sorella, and C. Filippi, J. Chem. Phys. **132**, 154113 (2010)

[57] W. J. Stevens, H. Basch, and M. Krauss, J. Chem. Phys. **81**, 6026 (1984)

[58] P. L. Rios, A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs, Phys. Rev. E **74**, 066701 (2006)

[59] J. Shepherd, G. Booth, A. Grüneis, and A. Alavi, Phys. Rev. B **85**, 081103(R) (2012)

[60] M. P. Nightingale and H. W. J. Blöte, Phys. Rev. Lett. **60**, 1562 (1988)

[61] N. Trivedi and D. M. Ceperley, Phys. Rev. B **40**, 2737 (1989); N. Trivedi and D. M. Ceperley, Phys. Rev. B **41**, 4552 (1990)

[62] H. J. M. van Bemmel, D. F. B. ten Haaf, W. van Saarloos, J. M. J. van Leeuwen, and G. An, Phys. Rev. Lett. **72**, 2442 (1994); D. F. B. ten Haaf, H. J. M. van Bemmel, J. M. J. van Leeuwen, and W. van Saarloos, Phys. Rev. B **51**, 13039 (1995)

[63] S. Sorella, Phys. Rev. Lett. **80**, 4558 (1998)

[64] G. H. Booth, A. J. W. Thom, and A. Alavi, J. Chem. Phys. **131**, 054106 (2009)

[65] D. Cleland, G. H. Booth, and A. Alavi, J. Chem. Phys. **132**, 041103 (2010)

[66] D. E. Woon and T. Dunning Jr, J. Chem. Phys. **98**, 1358 (1993)

[67] C. J. Umrigar, Phys. Rev. Lett. **71**, 408 (1993)