

Combining Orbital and Real Space Quantum Monte Carlo Methods

Von der Fakultät Chemie der Universität Stuttgart zur Erlangung der
Würde eines Doktors der Naturwissenschaften (Dr. rer. nat.)
genehmigte Abhandlung

Vorgelegt von
Seyed Mohammadreza Hosseini
aus Teheran, Iran

Hauptberichter: Prof. Dr. Ali Alavi
Mitberichter: Prof. Dr. Andreas Köhn
Prüfungsvorsitzender: Prof. Dr. Blazej Grabowski

Tag der mündlichen Prüfung: 18. April 2024

Institut für Theoretische Chemie der Universität Stuttgart

2024

Supervisor: Prof. Dr. Ali Alavi

Day-to-day Supervisor: Dr. Pablo López Ríos

Zusammenfassung

Quanten-Monte-Carlo (QMC)-Methoden sind wellenfunktionsbasierte Ansätze, die auf Zufallsstichproben beruhen. Diese Methoden bieten eine zuverlässige Lösung für Mehrkörperprobleme und liefern einen Vergleichsmaßstab (Benchmark), an den andere Methoden beurteilt werden können. Diese Arbeit beschäftigt sich mit der Entwicklung von QMC-Methoden zur Berechnung von hochpräzisen elektronischen Molekülenergien. Der Fokus richtet sich auf die Kombination von QMC-Verfahren, insbesondere Kontinuumsverfahren wie *Variational* und *Diffusion Quantum Monte Carlo* (VMC und DMC) mit *Full Configuration Interaction Quantum Monte Carlo* (FCIQMC). Das primäre Ziel dieser Arbeit ist es zu zeigen, dass diese Methoden aufgrund ihrer Komplementarität kombiniert werden können, um die jeweiligen Schwächen der einzelnen Methoden zu überwinden.

Einleitend wird eine Übersicht über die einzelnen QMC-Methoden auf der Grundlage der Literatur dargestellt. Nachfolgend werden drei Methoden untersucht, die für verschiedene Systemtypen anwendbar sind. Die erste Methode stellt eine neue Technik vor, mit der Gesamt- und Relativenergien sehr präzise durch Extrapolation von VMC- und DMC-Energien unter Verwendung von FCIQMC-generierten experimentellen Wellenfunktionen in Kombination mit einem Jastrow-Faktor und einer optionalen *Backflow*-Transformation erhalten werden können. Wir haben festgestellt, dass die VMC- und DMC-Energien glatte Funktionen der Summe der quadrierten Koeffizienten der FCIQMC-Wellenfunktion sind und, dass sich quadratische Extrapolationen der VMC- und DMC-Energien innerhalb der Unschärfe der exakten Gesamtenergie schneiden. Durch eine geeignete statistische Behandlung der quasi-zufälligen Fluktuationen liefert die *Extrapolated Intersection with Polynomials of Order Two* (XSPOT)-Methode Ergebnisse mit Benchmark-Qualität für die Gesamt- und Relativenergien verschiedener molekularer Systeme.

Auf Grundlage der Ergebnisse kombinierten wir die DMC-Methode und die *transcorrelated* FCIQMC (TC-FCIQMC)-Methode mit einem finiten Extrapolationsrahmen, um die thermodynamische Grenze der exakten Korrelationsenergie des hochdichten spin-unpolarisierten gleichförmigen Elektronengases zu berechnen. Die thermodynamische Grenze, die bei $r_s = 0,5$ ermittelt wurde, liegt innerhalb der Unschärfe der neuesten QMC-Korrelationsenergie aus der Literatur.

Im letzten Abschnitt unserer Arbeit stellen wir eine neue Methode vor, die ein neues Extrapolationsschema verwendet, das die Energieverzerrung eliminiert, die aus der *Initiator*-Approximation in der FCIQMC-Methode resultiert. Wir konnten ermitteln, dass die Extrapolation der FCIQMC-Energie linear in $n_w^{-1/3}$ zur *Infinite-Walker*-Grenze Ergebnisse liefert, die hervorragend mit dem Benchmark der *Full-Configuration-Interaction* (FCI)-Grenzwerte der semistochastischen *Heat-Bath-Configuration-Interaction* (SHCI)-Methode übereinstimmen. Die Methode wurde auf eine Reihe zweiatomiger Moleküle angewandt. Die Ergebnisse erwiesen sich als präzise, wobei die mit FCIQMC ermittelten FCI-Grenzwerte stets um 0,4 mHa oder weniger von den entsprechenden SHCI-Grenzwerten abwichen. Auch die Untersuchung des komplexen Chromatoms und seines Dimers war mit dieser Methode erfolgreich.

Abstract

Quantum Monte Carlo (QMC) methods are wave-function-based approaches based on random sampling that provide a very direct treatment of many-body problems and serve as benchmarks against which other techniques may be compared. This thesis is concerned with the development of QMC methods for the calculation of highly-accurate electronic energies. In particular, it focusses on combining QMC methods, namely continuum methods such as variational and diffusion Monte Carlo (VMC and DMC) with full configuration interaction quantum Monte Carlo (FCIQMC). The principal aim of this thesis is to show that thanks to their complementary nature, these methods can be used in combination to overcome specific weaknesses that each methodology has by itself.

The thesis begins with introductory chapters which describe the various QMC methods as they are used in the literature. Following this, three different methodologies are explored, which have applicabilities to different types of systems. In the first method, we present a new technique to obtain very accurate total and relative energies by extrapolating VMC and DMC energies, with the help of FCIQMC-generated trial wave functions combined with a Jastrow factor and an optional backflow transformation. We find that the VMC and DMC energies are smooth functions of the sum of the squared coefficients in the FCIQMC wave function, and that quadratic extrapolations of the VMC and DMC energies intersect within uncertainty of the exact total energy. With adequate statistical treatment of quasi-random fluctuations, the extrapolated intersection with polynomials of order two (XSPOT) method is shown to yield results in agreement with benchmark-quality total and relative energies for various molecular systems.

Following this, we combine the DMC and transcorrelated FCIQMC (TC-FCIQMC) methods, along with a finite-size extrapolation framework to calculate the thermodynamic limit of the exact correlation energy of the high-density spin-unpolarized uniform electron gas. The thermodynamic limit obtained at $r_s = 0.5$ is within uncertainty of the most recent QMC correlation energy from the literature.

In the final method, we introduce a new extrapolation scheme which eliminates the energy bias stemming from the initiator approximation in the FCIQMC method. We find that extrapolating the FCIQMC energy linearly in $n_w^{-1/3}$ to

the infinite-walker limit yields results in excellent agreement with the benchmark full configuration interaction (FCI) limits from the semistochastic heat-bath configuration interaction (SHCI) method. The method is applied to a series of diatomics and the results are accurate, with FCI limits obtained from FCIQMC differing by 0.4 mHa or less from their SHCI counterparts throughout. The more complex chromium atom and dimer are also successfully studied using this method.

Acknowledgements

“No one who achieves success does so without acknowledging the help of others. The wise and confident acknowledge this help with gratitude.”

—Alfred North Whitehead

The completion of this thesis is a momentous task and I would like to express my appreciation to everyone who has played a role. First and foremost, I owe a tremendous debt of gratitude to my supervisor, Prof. Dr. Ali Alavi, who has played a crucial role in my success over the past few years. His mentorship has been instrumental in shaping both the direction and depth of this thesis. I feel incredibly fortunate to have had the guidance and mentorship of such a highly respected academic. I am also grateful to Dr. Pablo López Ríos, whose constant support and constructive feedback have been indispensable in the completion of this research. His commitment to my progress and willingness to invest time and effort in my development has been invaluable.

I am particularly grateful to all my colleagues at Max Planck Institute for Solid State Research for their camaraderie, encouragement, and intellectual exchange throughout this journey. Especially, I want to thank Prof. Dr. David Tew for helpful suggestions, Dr. Hongjun Luo and Dr. Ke Liao for their help with the TC-FCIQMC calculations of the UEG, Dr. Werner Dobrautz, Dr. Kai Guthier, and Oskar Weser for helping with NECI code, Dr. Pradipta Samanta for carrying out multi-reference coupled cluster calculations, Dr. Daniel Kats for helping to generate CCSD natural orbitals with MOLPRO code, Dr. Michele Ruggeri for giving advice to twist average the FCIQMC energies, Dr. Alexander Lozovoi for providing FCIDUMP files for Cr atom and dimer, Dr. Vamshi Katukuri for the help with his extrapolation technique, and Dr. Khaldoun Ghanem for helping with the adaptive shift method.

For the proofreading of this thesis, I am grateful to Dr. Pablo López Ríos, Shiva Nosrati, and Mohammadreza Hasanlu. Especially, I thank Shiva for correcting the German abstract.

Living far away from my family has its challenges, but despite the physical distance, their love and encouragement has bridged the gap and given me a sense of connection and strength. My heartfelt thanks also go to my parents, sisters and friends for their support, understanding and encouragement through the ups and downs of the PhD process. Their love and support have been my anchor through it all. To my little niece, whose boundless energy brought joy and inspiration to my days and reminded me of the importance of curiosity and playfulness amidst the rigors of academia. I would also like to express

my heartfelt gratitude to my Shiva for her unwavering love and understanding throughout this journey. Her patience, support, and belief in me have been a constant source of strength and motivation, especially during moments of homesickness and longing for family.

To all those whose names are not explicitly mentioned but have contributed in various ways, whether through conversation, shared experiences, or acts of kindness, I extend my sincere thanks. Your support has played an integral role in my academic and personal growth, and I am deeply grateful for the collective encouragement and support I have received along the way.

Table of Contents

List of Tables	xiii
List of Figures	xv
I Introduction	1
II Theory	7
1 The Electronic Structure Problem	9
1.1 Atomic Units and Many-body Schrödinger Equation	9
1.2 The Born-Oppenheimer Approximation	11
1.3 Molecular Orbitals and Basis Sets	13
1.4 The Antisymmetry Principle, Slater Determinants and Configuration State Functions	14
1.5 Second Quantization	17
1.6 The Electronic Problem in Periodic Solids	18
1.6.1 The Bravais Lattice	19
1.6.2 Three Dimensional Uniform Electron Gas	20
1.6.3 Twist Averaging	21

2	Quantum Chemistry Approaches	23
2.1	Hartree-Fock Theory	23
2.2	Configuration Interaction	26
2.2.1	Heat-Bath Configuration Interaction	30
3	Quantum Monte Carlo Approaches	33
3.1	The Metropolis Algorithm	33
3.2	Variational Monte Carlo	34
3.3	Diffusion Monte Carlo	37
3.3.1	The Imaginary-time Schrödinger Equation	37
3.3.2	The Fixed-node Approximation	39
3.3.3	Importance Sampling	41
3.4	Trial Wave Functions	46
3.4.1	Slater-Jastrow Wave Functions	46
3.4.2	Multi-Slater-Jastrow Wave Functions	49
3.4.3	Backflow Wave Functions	50
3.5	Full Configuration Interaction Quantum Monte Carlo	51
3.5.1	The Transcorrelated FCIQMC	54
III	Results and Applications	61
4	Accurate Energies from Intersecting Extrapolations of Continuum Quantum Monte Carlo Data	63
4.1	Introduction	63
4.2	Methodology	65
4.2.1	Theoretical Justification	67
4.2.2	Obtaining Statistically Meaningful Results	69
4.2.3	Choice of Orbital Basis	74
4.3	Results and Discussion	74

5	Correlation Energy of the High-density Spin-unpolarized Electron Gas at the Thermodynamic Limit	81
5.1	Introduction	81
5.2	Methodology	83
5.2.1	Twist Averaging and Brillouin Zone Division	83
5.2.2	Basis-set Extrapolation of TC-FCIQMC Energies	85
5.2.3	Finite-size Extrapolation Framework for Fixed-node DMC	86
5.3	Results and Discussion	87
6	Walker Number Extrapolation of FCIQMC Energies	93
6.1	Introduction	93
6.2	Methodology	95
6.2.1	Modeling the Initiator Approximation	96
6.2.2	Scalar Measure of the Magnitude of the Initiator Perturbation	97
6.3	Results and Discussion	100
IV	Summary and Outlook	107
	Appendix A. The Slater-Condon Rules	111
	References	113

List of Tables

1.1	Conversion of atomic units to SI units.	10
4.1	Atoms and molecules considered in this work, along with their electronic states and geometries.	68
4.2	Fraction of spatial orbitals in the Hartree-Fock orbital basis expanded in each of four basis sets that appear in the first 300 configuration state functions of the full CI wave function for the ground state of the carbon dimer. . . .	69
4.3	Location of all six pairwise intersections of the VMC, DMC, bVMC, and bDMC curves shown in Figure 4.5 for the C_2 molecule. “Missed intersections” refer to random instances of the curves that do not intersect at $w > 1$ in the Monte Carlo resampling procedure.	73
4.4	Total energies in Ha, obtained with the XSPOT method using the VMC and bDMC data for the various atoms and molecules considered in this work, along with results from prior multi-determinant cQMC studies, our best individual bDMC energy for each system, and benchmark-quality non-relativistic total energies from the literature.	76
4.5	Atomization and excitation energies in mHa of the various molecules considered in this work corresponding to the total energies in Table 4.4 obtained from the XSPOT method, along with benchmark-quality non-relativistic relative energies from the literature.	76
5.1	Twist-averaged DMC correlation energies of the unpolarized UEG at $r_s = 0.5, 1.0$, and 2.0 and their respective thermodynamic limits, in mHa.	88
5.2	Twist-averaged TC-FCIQMC (“exact”) correlation energies and fixed-node errors for the unpolarized UEG at $r_s = 0.5$ for different system sizes and corresponding thermodynamic limits, in mHa.	91

5.3	Correlation energy of the unpolarized UEG at $r_s = 0.5$ obtained in this work, along with results from the VWN80, PZ81, PW92, and DPI parametrizations, and prior DMC studies, in mHa.	91
6.1	Electronic ground states and equilibrium bond lengths used for the first-row dimers considered in this work, following Ref. 1.	101
6.2	Total energies of relevant dimers in the FCI limit in Ha, obtained by extrapolation of SHCI energies to the zero second-order PT correction limit and by extrapolation of FCIQMC energies to the infinite-walker limit.	102
6.3	Total energies of the chromium atom and dimer, and atomization energies of the chromium dimer, in Ha. Results labeled “SHCI” are from Ref. 2, those labeled “FCIQMC” are large-population FCIQMC results obtained using natural orbitals from Ref. 3, and those labeled “This work” are FCIQMC energies obtained with Hartree-Fock orbitals extrapolated to the infinite-walker limit.	105

List of Figures

1.1	(a) 2D square Bravais lattice and (b) corresponding reciprocal lattice, including the first Brillouin zone (shaded in blue) and irreducible Brillouin zone (shaded in red).	20
2.1	CBS extrapolation of the Hartree-Fock energy of a carbon dimer C_2 at equilibrium geometry.	27
2.2	A singly-excited determinant $ \Psi_a^r\rangle$	28
2.3	Variation of CI energy of a carbon dimer C_2 at equilibrium geometry using the cc-pCVTZ basis with the number of determinants, including the estimated FCI energy from an FCIQMC calculation.	29
3.1	Flow chart of the VMC algorithm.	36
3.2	A 2-dimensional slice through the 321-dimensional nodal surface of a 2-dimensional Fermi gas containing 161 (polarized) fermions (reproduced with permission from Springer Nature. © Plenum Publishing Corporation 1991 ⁴). All 161 particle positions were sampled using VMC. The filled circle indicates the original position of the first particle. The other 160 particles are fixed at positions indicated by the open circles. The nodes of the wave function as a function of the position of the first particle are plotted.	40
3.3	Extrapolation of the DMC energy with respect to the time step for a carbon dimer C_2 at equilibrium geometry using a Slater-Jastrow trial wave function. The curve is a linear fit of the data. The width of the line reflects the statistical uncertainty of the fit values at each point.	43

3.4	DMC data for a Nitrogen dimer N_2 at equilibrium geometry with a Jastrow-free multi-determinant wave function. The upper panel shows the fluctuations in the population of walkers for the branching process arising from simulations done via Equation 3.30. In the lower panel, the black line shows the local energy averaged over the current population of walkers, the red line shows the reference energy, and the green line is the best estimate of the DMC energy. The initial walkers come from a previous VMC calculation, and the energy decays from its initial VMC value to reach a plateau. The data before the plateau correspond to the DMC equilibration phase.	45
3.5	Flow chart of the fixed-node DMC algorithm.	47
3.6	Local energy of a Nitrogen dimer N_2 at equilibrium geometry from the DMC calculation of a multi-determinant trial wave function (a) with and (b) without a Jastrow factor.	48
3.7	The upper plot shows the walker growth for a carbon dimer C_2 at equilibrium geometry in a cc-pCVDZ basis set in the fixed-shift mode ($S = 0$). After reaching a target population of 10^7 walkers, the shift, in the lower plot, starts to vary to keep the number of walkers constant. The projected energy is also shown in the lower plot. The shift value which stabilizes the walker growth is in agreement with the projected energy.	55
3.8	Finite-basis-set error in the FCIQMC calculations for the correlation energy of a carbon dimer C_2 at equilibrium geometry as a function of the inverse basis-set size M^{-1} . The data correspond to calculations in Dunning basis sets, cc-pCVnZ ($n=D, T, Q$). The line width indicates the statistical uncertainty in the fit.	56
3.9	Total energy of an unpolarized 14-electron gas at $r_s = 0.5$ obtained with FCIQMC and TC-FCIQMC methods as a function of the number of walkers.	59
4.1	VMC and DMC total ground-state energy of the carbon dimer using multi-determinant-Jastrow trial wave functions as a function of w , using Hartree-Fock orbitals expanded in the cc-pCVTZ basis set. Quadratic fits to the data are extended beyond $w = 1$ to show their intersection, which is in good agreement with the estimated exact non-relativistic total energy of the system. ⁵	67

4.2	Illustration of the expected behavior of cQMC energies as a function of w . The cQMC energies (circles and squares) deviate from the underlying smooth trend (lines) by a quasi-random amount (of amplitude represented by the width of the shaded area around the lines). The smooth trend can be represented by a quadratic function, $E(w)$ (dashed line), for $w_{\text{h.o.}} < w < w_{\text{exh.}}$ (shaded middle region), while for $w < w_{\text{h.o.}}$ higher-order contributions become important, and for $w > w_{\text{exh.}}$ orbital-basis exhaustion sets in. At the value of w_0 corresponding to the exact wave function the quadratic function gives the exact energy, $E_0 = E(w_0)$	70
4.3	Illustration of the Monte Carlo resampling scheme used to compute statistics on the intersection between two curves. For each curve, having obtained an estimate of α (width of the shaded region) from the original energy data (not shown), we create a synthetic instance of the dataset by shifting the original points by a random amount proportional to $\sqrt{\alpha^2 + \Delta E_i^2}$ (squares of same color saturation), perform a quadratic fit (line), and find the intersection between both fits (circled diamond). This process is repeated over the random instances (three shown in the illustration), from which statistics on the intersection are obtained.	71
4.4	VMC and DMC energies of the ground-state C_2 molecule as a function of w , as shown in Figure 4.1, using the full statistical treatment of the XSPOT method. The translucent areas represent the uncertainty in the values of the fit at each point. Also shown is the estimated exact non-relativistic energy ⁵ as a dotted line with a shaded area of ± 1 kcal/mol around it, and the intersection point between the curves. The inset shows the statistical distribution of intersection points as a color map with overlaid contour curves.	72
4.5	VMC and DMC energies of the ground-state C_2 molecule as a function of w , without (“VMC” and “DMC”) and with (“bVMC” and “bDMC”) a backflow transformation. The translucent areas represent the uncertainty in the values of the fit at each point. Also shown is the estimated exact non-relativistic energy ⁵ as a dotted line with a shaded area of ± 1 kcal/mol around it, and the intersection points between each of the six possible pairs of curves; error bars on these are only shown in the inset for clarity. . . .	73

4.6	VMC and DMC energies of the H ₂ O molecule as a function of w , both using Hartree-Fock orbitals (left) and CCSD natural orbitals (right) expanded in the cc-pCVTZ basis set. In both cases the same numbers of CSFs are used. The translucent areas represent the uncertainty in the values of the fit at each point. Also shown is the estimated exact non-relativistic energy ¹ as a dotted line with a shaded area of ± 1 kcal/mol around it, and the intersection point between the VMC and DMC curves in the left panel. The cQMC energies obtained with natural orbitals plateau with w , preventing the quadratic extrapolations from reaching the exact energy.	74
4.7	VMC and DMC energies of the atoms and molecules considered in this work as a function of w , without (“VMC” and “DMC”) and with (“bVMC” and “bDMC”) a backflow transformation. The translucent areas represent the uncertainty in the values of the fit at each point. Also shown in each plot is the relevant benchmark energy (see details in text and Table 4.4) as a dotted line with a shaded area of ± 1 kcal/mol around it, and the intersection point between the VMC and bDMC curves.	78
4.8	Number of distinct spatial orbitals in the FCIQMC trial wave functions relative to that in the orbital basis as a function of the number of CSFs, for the various systems considered in this work.	79
5.1	Finite basis-set error in the TC-FCIQMC correlation energy as a function of inverse basis-set size M^{-1} for the unpolarized 38-UEG at $r_s = 0.5$. The data correspond to calculations at multiple basis-set sizes in each of the IBZ regions. The translucent areas represent the uncertainty in the values of the fit at each point.	86
5.2	Finite-size corrected fixed-node correlation energies of the unpolarized UEG at $r_s = 0.5$, 1.0, and 2.0 as a function of N^{-1} . Horizontal dashed lines are shown at the thermodynamic limits for each density. The uncertainty in the fitted curves is smaller than the line width.	89
5.3	Fixed-node error for the unpolarized UEG at $r_s = 0.5$ as a function of N^{-1} . The linewidth represents the statistical uncertainty in the fit.	90
6.1	Estimated initiator error in the FCIQMC energies as a function of $n_w^{-1/3}$ for the CN molecule. The hollow points are not used in the extrapolations. The translucent areas represent the uncertainty in the values of the fit at each point.	101

6.2	Extrapolation to the FCI limit of SHCI energies with respect to the second-order PT correction E2 (left), and of FCIQMC energies with respect to $n_w^{-1/3}$ (right) for the CN molecule. The grayed-out points are not used in the FCIQMC extrapolations. The width of the line reflects the statistical uncertainty within the fit.	103
6.3	Estimated initiator error in the FCIQMC energies as a function of $n_w^{-1/3}$ for the chromium atom and dimer with the X2C relativistic Hamiltonian. The empty symbols denote points not used in the extrapolations. The line width represents the statistical uncertainty on the fit values at each point.	104

Part I

Introduction

Ab initio or first principle methods are numerical approaches to solve the many-body Schrödinger equation. They rely on established laws of nature and do not include empirical or semi-empirical parameters. The development of accurate *ab initio* methods remains a topical research subject of high interest. The three most popular families of *ab initio* methods are density functional theory (DFT)^{6,7}, quantum chemistry (QC)^{8,9}, and quantum Monte Carlo (QMC)^{10–14}. All these approaches have different regimes of applicability and different trade-offs between accuracy and computational complexity. QMC algorithms need little memory and are almost embarrassingly parallel, which makes them ideally suited to modern supercomputers. There exist numerically exact QMC methods to solve the many-body problem for bosonic systems. For fermionic systems, there also exist very good approximate, yet typically very accurate, QMC methods such as variational Monte Carlo (VMC)^{10,14}, fixed-node diffusion Monte Carlo (DMC)^{11,14}, and full configuration interaction quantum Monte Carlo (FCIQMC)^{12,13} methods.

The VMC method combines the variational principle with the Monte Carlo evaluation of integrals in real space. The expectation values of operators such as the Hamiltonian are calculated in VMC with an approximate many-body trial wave function, whose parameters are optimized during the calculation. In general, there are no restrictions on the form of the trial wave function. A Jastrow factor^{15,16} multiplied by a Slater determinant is the most commonly used form of a trial wave function in QMC. To get higher levels of accuracy however, a more sophisticated ansatz must be used. This is particularly the case in intrinsically multi-reference problems. A Jastrow factor multiplied by a truncated configuration interaction (CI) expansion is also a popular and systematically improvable trial wave function. Such an ansatz combined with a backflow function¹⁷ constitutes the most complex trial wave function used in this thesis.

The DMC method is a real-space projection technique in which higher accuracy is achieved by evolving the wave function in imaginary time towards the lowest-energy solution of the Schrödinger equation. The fixed-node approximation is used to maintain fermionic antisymmetry so that the intricate nodal surface of the DMC wave function is constrained to equal that of the approximate trial wave function optimized in VMC. This results in a positive bias in the computed energy, known as the fixed-node error. The VMC and DMC methods are categorized as continuum QMC (cQMC) methods; the DMC energy is bounded from above by the VMC energy and from below by the exact energy.

Full configuration interaction (FCI) stands out as one of the most robust techniques for addressing electronic correlations in many-body systems. However, substantial computational cost and memory requirements associated with iterative diagonalization within the Slater determinant space significantly constrains its practical implementation. The FCIQMC method is a second-quantized projection technique that converges onto the FCI

solution by randomly sampling Slater-determinant space. The FCIQMC method not only requires significantly less memory than an FCI calculation, but unlike the iterative diagonalization procedure required in FCI, the FCIQMC algorithm is readily parallelizable and can be run efficiently on thousands of computer cores. Integration in Slater-determinant space also reduces the sign problem in fermionic systems^{18,19}. This method is based on simulating the population dynamics of a set of walkers, such that in the limit of large walker numbers, the population effectively samples the exact fermionic ground-state wave function. However, the required number of walkers is proportional to the size of the full Hilbert space, making the method unfeasible for large systems. The initiator approximation¹³ substantially reduces this number by introducing a systematic bias, known as the initiator error.

The synergistic strengths of cQMC and FCIQMC motivate us to develop methods combining these approaches. As the first combined methodology, we present a novel extrapolation scheme for the VMC and DMC results that yields very accurate total and relative energies. The FCIQMC method is used to generate multi-determinant wave functions, which are optimized in the presence of a Jastrow factor and an optional backflow transformation in the VMC method. The optimized trial wave functions are then employed in the DMC calculations. In the end, we find that the VMC and DMC energies are smooth functions of the sum of the squared coefficients of the initial wave function, and that pairwise intersections of quadratic fits to the VMC and DMC energies lie within uncertainty of the exact total energy. Using modest CI expansion sizes, the extrapolated intersection with polynomials of order two (XSPOT) method is demonstrated to yield results aligned with benchmark-quality total and relative energies.

The understanding of electronic correlations in real materials crucially involves the study of the three-dimensional uniform electron gas (UEG)²⁰, since its correlation energy as a function of density is used in DFT calculations to describe the behavior of electrons. In this thesis, we make use of the finite-size extrapolation framework proposed in Ref. 21 to estimate the thermodynamic limit of the exact correlation energy of the high-density spin-unpolarized UEG. We treat the transcorrelated FCIQMC energy as exact and then we calculate its thermodynamic limit with the help of the finite-size extrapolation of the fixed-node DMC energy and fixed-node error. The obtained thermodynamic limit at $r_s = 0.5$ lies within uncertainty of the most recent correlation energy from the literature.

To eliminate the initiator bias in FCIQMC calculations, we develop a new model for the initiator approximation to extrapolate the energy bias. We find that extrapolating the FCIQMC total energies calculated at different walker numbers n_w linearly in $n_w^{-1/3}$ to the infinite-walker limit removes the initiator error and converges the energy to the FCI limit. The application of this method to a series of dimers gives results in excellent agreement

with FCI limits estimated with the semistochastic heat-bath CI (SHCI)²² method.

This thesis is structured as follows. Part II contains some of the basic concepts, tools and theories relevant to this thesis. We introduce the electronic structure problem in some detail. Starting from the many-body Schrödinger equation, the Born-Oppenheimer approximation, molecular orbitals, the antisymmetry principle, and second quantization are then discussed. The electronic problem in periodic solids is explored by introducing the Bravais lattice, the three-dimensional uniform electron gas, and twist averaging. Then, we discuss the two most important quantum chemistry approaches employed in QMC calculations, namely Hartree-Fock theory and configuration interaction. The QMC methods including variational and diffusion Monte Carlo and full configuration interaction quantum Monte Carlo are also detailed in the last section. Part III focuses on the new theoretical developments and applications, where the three different methods are presented. In Part IV, we conclude the thesis by summarizing the main contents and offering insights into potential future directions.

Part II

Theory

Chapter 1

The Electronic Structure Problem

This chapter introduces the basic concepts, definitions, and notations in quantum chemistry. Here we aim to formulate the problem of describing the motion of electrons in the field of fixed nuclear point charges. Starting from the Born-Oppenheimer approximation, we then discuss the molecular orbitals, the antisymmetry principle, and second quantization. Some properties of periodic solids are then discussed in the final section ^{8,9}.

1.1 Atomic Units and Many-body Schrödinger Equation

The units we use throughout this thesis are called Hartree atomic units. Formally when doing computer calculations, it is preferable to choose a system of units in which the important physical quantities have numerical values of order of 1. If we leave factors of 6.63×10^{-34} and 1.6×10^{-19} floating around, we could run into overflow or underflow errors or rounding-error problems. With an appropriate choice of units, the numerical values of $\hbar$, e , m_e , and $4\pi\epsilon_0$ can all be made equal to 1. Consequently, the dimensionless Schrödinger equation then looks simpler.

The length and the energy scales in atomic units are the Bohr radius,

$$1 \text{ Bohr radius} = a_0 = \frac{4\pi\epsilon_0\hbar^2}{m_e e^2}, \quad (1.1)$$

and the Hartree,

$$1 \text{ Hartree} = E_a = \frac{\hbar^2}{m_e a_0^2} = \frac{e^2}{4\pi\epsilon_0 a_0} = \frac{me^4}{(4\pi\epsilon_0\hbar)^2}, \quad (1.2)$$

respectively. Table 1.1 shows the important physical quantities and conversion factors X between atomic units and SI units, such that the value of any quantity in SI units Q is related to that in atomic units Q' by $Q = XQ'$.

Physical quantity	Conversion factor X	Value of X (SI)
Length	a_0	$5.2918 \times 10^{-11} \text{ m}$
Mass	m_e	$9.1095 \times 10^{-31} \text{ kg}$
Charge	e	$1.6022 \times 10^{-19} \text{ C}$
Energy	E_a	$4.3598 \times 10^{-18} \text{ J}$
Angular Momentum	$\hbar$	$1.0546 \times 10^{-34} \text{ J s}$

Table 1.1: Conversion of atomic units to SI units.

In this thesis, we aim to find approximate solutions to the non-relativistic time-independent Schrödinger equation,

$$\hat{H}\Phi = E\Phi, \quad (1.3)$$

where $\hat{H}$ is the many-body Hamiltonian for a system of electrons and nuclei, and is defined as

$$\hat{H} = -\sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}} + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}}, \quad (1.4)$$

where M_A is the ratio of the mass of the nucleus A to the mass of an electron and Z_A is the atomic number of the nucleus A . Electrons and nuclei are described by position vectors $\mathbf{r}_i$ and $\mathbf{R}_A$, respectively. $r_{ij} = |\mathbf{r}_{ij}| = |\mathbf{r}_i - \mathbf{r}_j|$ is the distance between the i th and j th electron, $r_{iA} = |\mathbf{r}_i - \mathbf{R}_A|$ is the distance between the i th electron and A th nucleus, and $R_{AB} = |\mathbf{R}_A - \mathbf{R}_B|$ is the distance between the A th and B th nucleus. The first term in the formula is an operator for the kinetic energy of electrons; the second term is the kinetic energy for nuclei; the third one is the attractive Coulomb potential between electrons and nuclei, and the fourth and fifth terms are the repulsive Coulomb potentials between electrons and between nuclei, respectively.

1.2 The Born-Oppenheimer Approximation

In order to tackle the task of solving the Schrödinger equation, Equation 1.3, one typically employs the following separation ansatz,

$$\Phi(\mathbf{r}_i, \mathbf{R}_A) = \psi(\mathbf{r}_i)\phi(\mathbf{R}_A), \quad (1.5)$$

where the wave function is written as a product of the electronic $\psi(\mathbf{r}_i)$ and nuclear $\phi(\mathbf{R}_A)$ wave functions, and $\mathbf{r}_i$ and $\mathbf{R}_A$ are the set of electronic and nuclear coordinates, respectively. Unfortunately, the fact that the electron-nucleus interaction term (the third term in 1.4) depends on both electronic and nuclear coordinates prevents the exact separation of $\hat{H}$ into electronic and nuclear parts. The Born-Oppenheimer approximation^{23,24} is the assumption that the motion of electrons and nuclei in a molecule can be separated. Since the nuclei are so much heavier than electrons, one can consider, to a good approximation, that electrons move in the field of fixed nuclei. Then, in Equation 1.4 the nuclear kinetic energy term (the second term) can be neglected, and one can also ignore the nuclear potential energy term (the last term) as it is constant. The remaining terms in Equation 1.4 constitute the so-called electronic Hamiltonian, which describes the motion of N electrons in the field of M point charges,

$$\hat{H}_{elec} = - \sum_{i=1}^N \frac{1}{2} \nabla_i^2 - \sum_{i=1}^N \sum_{A=1}^M \frac{Z_A}{r_{iA}} + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{r_{ij}}, \quad (1.6)$$

where the corresponding electronic Schrödinger equation is written as

$$\hat{H}_{elec}\psi(\mathbf{r}_i, \mathbf{R}_A) = E_{elec}(\mathbf{R}_A)\psi(\mathbf{r}_i, \mathbf{R}_A), \quad (1.7)$$

in which the electronic wave function $\psi(\mathbf{r}_i, \mathbf{R}_A)$ depends explicitly on the electronic coordinates and parametrically on the nuclear coordinates, as does the electronic energy $E_{elec}(\mathbf{R}_A)$. Parametric dependence means that for different arrangements of the nuclei, ψ is a different function of the electronic coordinates. Solving the electronic Schrödinger equation produces a whole set of electronic wave functions $\{\psi_\alpha(\mathbf{r}_i, \mathbf{R}_A)\}$, which for any value of $\mathbf{R}_A$, can be considered to be orthonormal and complete. The total wave function $\Phi(\mathbf{r}_i, \mathbf{R}_A)$ can then be expanded in terms of the electronic basis functions,

$$\Phi(\mathbf{r}_i, \mathbf{R}_A) = \sum_{\alpha} \psi_{\alpha}(\mathbf{r}_i, \mathbf{R}_A)\phi_{\alpha}(\mathbf{R}_A), \quad (1.8)$$

where the nuclear wave function $\phi_\alpha(\mathbf{R}_A)$ describes the motion of the nuclei on the potential energy surface associated with the electronic state α . The above expansion is then inserted into Equation 1.3 to yield

$$\hat{H} \sum_{\alpha'} \psi_{\alpha'}(\mathbf{r}_i, \mathbf{R}_A) \phi_{\alpha'}(\mathbf{R}_A) = E \sum_{\alpha''} \psi_{\alpha''}(\mathbf{r}_i, \mathbf{R}_A) \phi_{\alpha''}(\mathbf{R}_A). \quad (1.9)$$

After substituting Equations 1.4 and 1.6 into 1.9, one can multiply it by $\psi_\alpha^*(\mathbf{r}_i, \mathbf{R}_A)$ from left and then integrate over $\mathbf{r}_i$ to yield

$$\begin{aligned} E\phi_\alpha(\mathbf{R}_A) &= \sum_{\alpha'} \left\langle \psi_\alpha(\mathbf{r}_i, \mathbf{R}_A) \left| \hat{H}_{elec} + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}} \right| \psi_{\alpha'}(\mathbf{r}_i, \mathbf{R}_A) \right\rangle \phi_{\alpha'}(\mathbf{R}_A) \\ &\quad - \sum_{\alpha'} \sum_{A=1}^M \frac{1}{2M_A} \langle \psi_\alpha(\mathbf{r}_i, \mathbf{R}_A) | \nabla_A^2 | \psi_{\alpha'}(\mathbf{r}_i, \mathbf{R}_A) \rangle \phi_{\alpha'}(\mathbf{R}_A), \end{aligned} \quad (1.10)$$

where the second term on the right-hand side can be expanded using the chain rule as

$$\begin{aligned} \langle \psi_\alpha(\mathbf{r}_i, \mathbf{R}_A) | \nabla_A^2 | \psi_{\alpha'}(\mathbf{r}_i, \mathbf{R}_A) \rangle \phi_{\alpha'}(\mathbf{R}_A) &= \delta_{\alpha\alpha'} \nabla_A^2 \phi_{\alpha'}(\mathbf{R}_A) \\ &\quad + 2 \langle \psi_\alpha(\mathbf{r}_i, \mathbf{R}_A) | \nabla_A | \psi_{\alpha'}(\mathbf{r}_i, \mathbf{R}_A) \rangle \cdot \nabla_A \phi_{\alpha'}(\mathbf{R}_A) \\ &\quad + \phi_{\alpha'}(\mathbf{R}_A) \langle \psi_\alpha(\mathbf{r}_i, \mathbf{R}_A) | \nabla_A^2 | \psi_{\alpha'}(\mathbf{r}_i, \mathbf{R}_A) \rangle. \end{aligned} \quad (1.11)$$

Substitution of Equation 1.11 into 1.10 results in a set of coupled equations for $\phi_\alpha(\mathbf{R}_A)$,

$$\begin{aligned} E\phi_\alpha(\mathbf{R}_A) &= E'_{elec,\alpha}(\mathbf{R}_A) \phi_\alpha(\mathbf{R}_A) - \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 \phi_\alpha(\mathbf{R}_A) \\ &\quad - \sum_{\alpha'} \sum_{A=1}^M \frac{1}{2M_A} \left[2\mathbf{F}_{\alpha\alpha'}(\mathbf{R}_A) \cdot \nabla_A \phi_{\alpha'}(\mathbf{R}_A) + \phi_{\alpha'}(\mathbf{R}_A) G_{\alpha\alpha'}(\mathbf{R}_A) \right], \end{aligned} \quad (1.12)$$

where

$$E'_{elec,\alpha}(\mathbf{R}_A) = E_{elec,\alpha}(\mathbf{R}_A) + \sum_{A=1}^M \sum_{B>A}^M \frac{Z_A Z_B}{R_{AB}}, \quad (1.13)$$

are the potential energy surfaces that govern the nuclear motion. The additional coupling terms

$$\mathbf{F}_{\alpha\alpha'}(\mathbf{R}_A) = \langle \psi_\alpha(\mathbf{r}_i, \mathbf{R}_A) | \nabla_A | \psi_{\alpha'}(\mathbf{r}_i, \mathbf{R}_A) \rangle, \quad (1.14)$$

$$\mathbf{G}_{\alpha\alpha'}(\mathbf{R}_A) = \langle \psi_\alpha(\mathbf{r}_i, \mathbf{R}_A) | \nabla_A^2 | \psi_{\alpha'}(\mathbf{r}_i, \mathbf{R}_A) \rangle, \quad (1.15)$$

are called the non-adiabatic interactions, which are responsible for transitions between potential energy surfaces. These couplings are typically small. If they can be safely neglected, and assuming that the wave function is real, one will end up with

$$\left[- \sum_{A=1}^M \frac{1}{2M_A} \nabla_A^2 + E'_{elec,\alpha}(\mathbf{R}_A) \right] \phi_\alpha(\mathbf{R}_A) = E \phi_\alpha(\mathbf{R}_A), \quad (1.16)$$

which describes for the motion of the nuclei on a given Born-Oppenheimer potential energy surface obtained by solving the electronic Schrödinger equation. Also, E is called the Born-Oppenheimer approximation to the total energy of Equation 1.3, which includes the electronic, vibrational, rotational, and translational energy.

For the rest of this thesis, we just concentrate on the electronic problem (Equations 1.6 and 1.7). We thus drop the subscript “*elec*” and work with the electronic Schrödinger equation.

1.3 Molecular Orbitals and Basis Sets

Molecular orbitals are defined as wave functions of individual, non-interacting electrons in a molecule. Spatial orbital $\psi_i(\mathbf{r})$ represents the spatial distribution of the non-interacting electron, such that $|\psi_i(\mathbf{r})|^2 d\mathbf{r}$ is the probability of finding the electron in the small volume of $d\mathbf{r}$ surrounding $\mathbf{r}$. Spatial orbitals are usually assumed to form an orthonormal set,

$$\int \psi_i^*(\mathbf{r}) \psi_j(\mathbf{r}) d\mathbf{r} = \delta_{ij}. \quad (1.17)$$

In practice, the set is not complete as this would require an infinite number of functions. A finite set $\{\psi_i | i = 1, 2, 3, \dots, K\}$ of K spatial orbitals will only cover a certain region of the complete space; however, we can always have “exact” results within a finite set. Taking the electron spin into account, we would need spin functions to describe up and down spins. A complete set for describing the spin of electrons consists of two orthonormal functions $\alpha(\omega)$ and $\beta(\omega)$ for spin up ($\uparrow$) and spin down ($\downarrow$), respectively. A wave function of an electron describing both its spatial distribution and its spin is called spin orbital $\chi(\mathbf{x})$, where $\mathbf{x}$

indicates both spatial and spin coordinates. Therefore, one can form two different spin orbitals from any spatial orbitals as follows,

$$\chi(\mathbf{x}) = \begin{cases} \psi(\mathbf{r})\alpha(\omega) \\ \text{or} \\ \psi(\mathbf{r})\beta(\omega) \end{cases}. \quad (1.18)$$

Then with a set of K orthonormal spatial orbitals one can construct $2K$ orthonormal spin orbitals,

$$\int \chi_i^*(\mathbf{x})\chi_j(\mathbf{x}) d\mathbf{x} = \delta_{ij}. \quad (1.19)$$

Spatial orbitals are also expressed as linear combinations of a set of functions, called basis functions. A basis set is then defined as a set of spatial basis functions $\{\phi_\mu(\mathbf{r})|\mu = 1, 2, \dots, K\}$ used to represent electronic wave function. A set of K basis functions would produce K spatial orbitals and $2K$ spin orbitals.

1.4 The Antisymmetry Principle, Slater Determinants and Configuration State Functions

Here we discuss the wave function for a collection of electrons, i.e., the N -electron wave function. So far we have seen that a spin orbital $\chi(\mathbf{x})$ can be considered as an appropriate wave function for a single electron. The wave function of an N -electron system then must be a function of $\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N$ as $\Phi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N)$. The antisymmetry principle is a general statement of the Pauli exclusion principle, and indicates that the many-fermionic wave function must be antisymmetric with respect to the interchange of the coordinate $\mathbf{x}$ of any two fermions. The exact wave function Φ in terms of this principle would then be as follows,

$$\Phi(\mathbf{x}_1, \dots, \mathbf{x}_i, \dots, \mathbf{x}_j, \dots, \mathbf{x}_N) = -\Phi(\mathbf{x}_1, \dots, \mathbf{x}_j, \dots, \mathbf{x}_i, \dots, \mathbf{x}_N). \quad (1.20)$$

Thus the N -electron wave function not only has to satisfy the Schrödinger equation, but it also must be antisymmetric as in Equation 1.20. An antisymmetric N -electron wave function based on the single-particle-orbital basis $\{\chi_i\}$ can be written in a determinant

form as

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = (N!)^{-1/2} \begin{vmatrix} \chi_i(\mathbf{x}_1) & \chi_j(\mathbf{x}_1) & \dots & \chi_k(\mathbf{x}_1) \\ \chi_i(\mathbf{x}_2) & \chi_j(\mathbf{x}_2) & \dots & \chi_k(\mathbf{x}_2) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_i(\mathbf{x}_N) & \chi_j(\mathbf{x}_N) & \dots & \chi_k(\mathbf{x}_N) \end{vmatrix}, \quad (1.21)$$

which is called Slater determinant. The factor $(N!)^{-1/2}$ serves as a normalization factor. The N -electron Slater determinant represents N electron occupying N spin orbital $(\chi_i, \chi_j, \dots, \chi_k)$, without specifying which electron occupies which spin orbital (electrons are indistinguishable). Each row of the determinant is labeled by an electron, and each column corresponds to a specific spin orbital. Interchanging any two rows of the determinant leaves a negative sign, which is in agreement with the antisymmetry principle. Besides, any two electrons occupying the same spin orbitals, any two equal rows, making the determinant zero. The last statement is basically nothing but the familiar Pauli exclusion principle, which states that no more than one electron can occupy a spin orbital. Moreover, it can be proved that this single determinantal approximation to the wave function provides a correlated motion of parallel-spin electrons. This effect is referred to as exchange correlation.

It is also convenient to use the short-hand notation

$$\Psi(\mathbf{x}_1, \mathbf{x}_2, \dots, \mathbf{x}_N) = |\chi_i \chi_j \dots \chi_k\rangle, \quad (1.22)$$

where the antisymmetry principle reads

$$|\dots \chi_i \dots \chi_j \dots\rangle = -|\dots \chi_j \dots \chi_i \dots\rangle. \quad (1.23)$$

Slater determinants formed from different orthonormal spin orbitals are orthonormal. Also, Slater determinants are generally not eigenstates of the $\hat{\mathbf{S}}^2$ operator, thus in general they are not pure spin states. However, one can construct a basis of $\hat{\mathbf{S}}^2$ eigenfunctions using techniques such as the unitary group approach (UGA)²⁵. In the the UGA framework, the N -electron spin-adapted Hilbert space is spanned by the orthonormal Gel'fand-Tsetlin (GT)^{26–28} basis. The constituent elements of this basis, denoted as configuration state functions (CSFs), are constructed from a set of n orthonormal spatial orbitals and can be described uniquely by a three-column table, known as Paldus tableau²⁵,

$$[p] = \begin{bmatrix} a_n & b_n & c_n \\ a_{n-1} & b_{n-1} & c_{n-1} \\ \vdots & \vdots & \vdots \\ a_1 & b_1 & c_1 \\ 0 & 0 & 0 \end{bmatrix}, \quad (1.24)$$

where the first column, a_i , contains the number of doubly-occupied orbitals, the second column, b_i , contains the number of singly-occupied orbitals, and the last one, c_i , contains the number of empty orbitals. Therefore, a_i , b_i , and c_i are restricted to be non-negative integers, while satisfying the condition $a_i + b_i + c_i = i$. Besides, if $\Delta x_i = x_i - x_{i-1}$ ($x = a, b, c$), then Δa_i and Δc_i are restricted to be 0 or 1, while Δb_i can be -1, 0, or 1, with $\Delta a_i + \Delta b_i + \Delta c_i = 1$. All possible CSFs of a GT basis can also be encoded in a vector named as the step vector $\mathbf{d} = (d_1, d_2, \dots, d_n)$. Similar to the Paldus tableau, this vector focuses on the occupancy and spin coupling of the individual orbitals, and its elements are defined as

$$d_i = 3\Delta a_i + \Delta b_i, \quad (1.25)$$

where d_i is allowed to take on four distinct values: an empty orbital is indicated by $d_i = 0$, a singly-occupied orbital with $\Delta S_i = 1/2$ by $d_i = 1$ and with $\Delta S_i = -1/2$ by $d_i = 2$, and a doubly-occupied orbital by $d_i = 3$.²⁹ The compilation of all d_i values into the step vector $\mathbf{d}$ format constitutes the most efficient way to represent a CSF, maintaining the same storage requirement as a Slater determinant, utilizing 2 bits for each spatial orbital.

CSFs can be expressed explicitly in terms of Slater determinants,

$$|\Psi_{\text{CSF}}\rangle = \sum_I c_I |D_I\rangle, \quad (1.26)$$

where $|\Psi_{\text{CSF}}\rangle$ and $|D_I\rangle$ are CSF- and Slater determinant-based wave functions, respectively, and c_I are coefficients. Once the M_S value, the eigenvalue of the spin component operator $\hat{S}_Z$, has been selected, Slater determinants can be constructed by attaching $(N/2 + M_S)$ α (up) spins and $(N/2 - M_S)$ β (down) spins to the occupied orbitals in all possible ways, ensuring that each doubly-occupied orbital has one occurrence associated with α spin and the other with β spin, following the specified sequence. Also, the coefficient of each Slater determinant in Equation 1.26 can be obtained as a product over spatial orbitals,

$$c = \prod_{i=1}^n f_i, \quad (1.27)$$

where the f_i values are given by

$$\begin{aligned} f_i &= 1 && \text{if } d_i = 0, \\ &= \sqrt{(a_i + b_i - \gamma_i)/b_i} && \text{if } d_i = 1, \\ &= (-1)^{b_i + \delta_i} \sqrt{(\gamma_i - a_i + 1)/(b_i + 2)} && \text{if } d_i = 2, \\ &= (-1)^{b_i} && \text{if } d_i = 3, \end{aligned} \quad (1.28)$$

where γ_i is the number of spin orbitals preceding orbital i in the Slater determinant which have spin opposite to that of orbital i , and δ_i is 1 or 0 if orbital i has spin α or β , respectively.²⁹

The total number of CSFs corresponding to a configuration of n orbitals, N electrons, and total spin S is determined using the Weyl-Paldus^{25,30} dimension formula,

$$N_{\text{CSF}} = \frac{2S+1}{n+1} \binom{n+1}{\frac{N}{2}-S} \binom{n+1}{n-\frac{N}{2}-S}. \quad (1.29)$$

The ratio between the total number of Slater determinants and CSFs for $N = n$, provided that n and N are sufficiently large, can be estimated using Stirling's formula as

$$\frac{N_{\text{SD}}}{N_{\text{CSF}}} \approx \frac{\sqrt{\pi n n}}{2(2S+1)}, \quad (1.30)$$

which demonstrates an orbital-dependent reduction in the size of the Hilbert space for a spin-adapted basis.

1.5 Second Quantization

The antisymmetry principle serves as a fundamental axiom in quantum mechanics, and is satisfied by employing Slater determinants and linear combinations of such determinants for wave functions. Second quantization is a formalism which satisfies the antisymmetry principle without using Slater determinants, where the wave function's antisymmetry property is transferred onto the algebraic properties of certain operators. In order to construct the formalism of second quantization, we first introduce the creation operator $\hat{a}_i^\dagger$, which creates an electron in the spin orbital χ_i as follows,

$$\hat{a}_i^\dagger |\chi_j \dots \chi_k\rangle = |\chi_i \chi_j \dots \chi_k\rangle. \quad (1.31)$$

The adjoint of the creation operator (i.e., $(\hat{a}_i^\dagger)^\dagger = \hat{a}_i$) is called the annihilation operator and is defined by

$$\hat{a}_i |\chi_i \chi_j \dots \chi_k\rangle = |\chi_j \dots \chi_k\rangle, \quad (1.32)$$

which annihilates or destroys an electron in spin orbital χ_i . All the properties of Slater determinants can be found in the anticommutation relations between two creation operators,

between two annihilation operators, and between a creation and an annihilation operator as follows,

$$\{\hat{a}_i^\dagger, \hat{a}_j^\dagger\} = 0, \quad \{\hat{a}_i, \hat{a}_j\} = 0, \quad \{\hat{a}_i, \hat{a}_j^\dagger\} = \delta_{ij}, \quad (1.33)$$

where the anticommutation relation is defined as $\{\hat{A}, \hat{B}\} = \hat{A}\hat{B} + \hat{B}\hat{A}$. Slater determinants can be represented within this formalism as

$$\hat{a}_i^\dagger \hat{a}_j^\dagger \dots \hat{a}_k^\dagger | \rangle = |\chi_i \chi_j \dots \chi_k\rangle, \quad (1.34)$$

where $| \rangle$ is called the vacuum state which represents a state of the system that contains no electrons. In order to develop the entire theory of many-body systems without determinants, one must express the many-body operators $\hat{\mathcal{O}}_1$ and $\hat{\mathcal{O}}_2$ (explained later in chapter 2 and Appendix A) in terms of creation and annihilation operators as

$$\hat{\mathcal{O}}_1 = \sum_{ij} [i|\hat{h}|j] \hat{a}_i^\dagger \hat{a}_j \quad (1.35)$$

$$\hat{\mathcal{O}}_2 = \frac{1}{2} \sum_{ijkl} [ik|jl] \hat{a}_i^\dagger \hat{a}_j^\dagger \hat{a}_l \hat{a}_k, \quad (1.36)$$

where the summations run over the set of spin orbitals, and $[i|\hat{h}|j]$ and $[ik|jl]$ are called one-body and two-body integrals, respectively (explained later in chapter 2). Note that any result that can be obtained using the properties of determinants can also be proved using only the algebra of creation and annihilation operators.

1.6 The Electronic Problem in Periodic Solids

So far we have investigated the electronic structure problem in molecular systems, where the boundary condition only implies that the wave function must vanish at infinity. In periodic systems, e.g., the uniform electron gas, simulation cells subjected to periodic boundary conditions are often used to account for the periodic arrangement of atoms in crystals. In order to accurately compare theoretical results with experimental findings, it is necessary to extrapolate all calculated observables to the thermodynamic limit (TDL). This involves simulating increasing numbers of particles and simulation-cell volumes, while maintaining a constant ratio between them,

$$N \rightarrow \infty, \quad V \rightarrow \infty, \quad \frac{N}{V} = \text{Constant}. \quad (1.37)$$

The ensuing subsections cover some fundamental aspects of the periodic systems' structure, including an introduction to the three-dimensional uniform electron gas and to twist averaging.

1.6.1 The Bravais Lattice

A crystal structure is formed by a set of atoms or groups of them which are periodically distributed. The starting point for the formal description of crystals is to introduce a mathematical abstraction called lattice, which is a periodic arrangement of points in space, whose positions are given by the translation lattice vector,

$$\mathbf{T} = n_1 \mathbf{a}_1 + n_2 \mathbf{a}_2 + n_3 \mathbf{a}_3 \quad n_i \in \mathcal{Z}, \quad (1.38)$$

where $\mathbf{a}_i$ are called primitive lattice vectors or basis vectors. The translation lattice vector can have any point in three-dimensional space as reference. The basis vectors $\mathbf{a}_i$ define the smallest volume of the lattice, called the primitive unit cell, that can be repeated in three directions to construct the entire lattice. The Wigner-Seitz unit cell is a primitive unit cell and refers to the volume of space surrounding a single point in the lattice that is closer to that point than any other point in the lattice. The choice of primitive lattice vectors and the primitive unit cell in a lattice is not unique, however all possible choices of primitive unit cells have the same volume associated with exactly one lattice point. Finally, when an atom or a set of atoms is added to every lattice point that lies inside the primitive unit cell volume, the so-called basis is created, resulting in the crystal structure.

According to their point symmetry, lattices are classified into different systems, such that all lattices that possess the same point symmetry belong to a particular lattice system. In this sense, there are in total 4 possible lattice systems including 5 different lattices, the so-called Bravais lattices, in two-dimensional space, and 7 possible lattice systems including 14 Bravais lattices in three-dimensional space. One can also define the reciprocal lattice which has the same point symmetry as the crystal lattice (direct lattice) and plays an important role in the description of the physical properties of crystalline materials. For a crystal lattice with the translation vector $\mathbf{T}$, the corresponding reciprocal lattice is defined as

$$e^{i\mathbf{T} \cdot \mathbf{G}} = 1, \quad \mathbf{a}_i \cdot \mathbf{b}_j = 2\pi\delta_{ij}, \quad (1.39)$$

where the basis vectors and translation vector of the reciprocal lattice are represented by $\mathbf{b}_j$ and $\mathbf{G}$, respectively, and satisfy the following relation,

$$\mathbf{G} = m_1 \mathbf{b}_1 + m_2 \mathbf{b}_2 + m_3 \mathbf{b}_3 \quad m_i \in \mathcal{Z}. \quad (1.40)$$

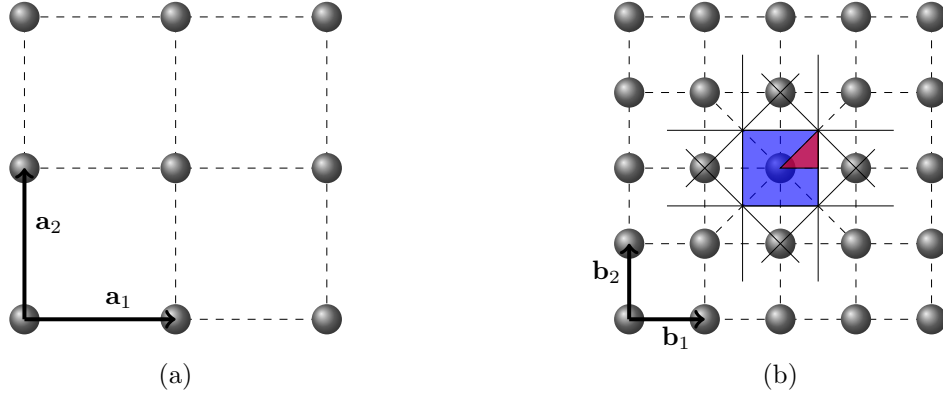


Figure 1.1: (a) 2D square Bravais lattice and (b) corresponding reciprocal lattice, including the first Brillouin zone (shaded in blue) and irreducible Brillouin zone (shaded in red).

The first Brillouin zone is the Wigner-Seitz unit cell constructed in the reciprocal lattice. The irreducible Brillouin zone (IBZ) is also defined as the smallest space volume within the Brillouin zone that completely describes the periodic structure of the lattice. Consequently, in the presence of additional symmetry, the irreducible Brillouin zone is smaller than the Brillouin zone. Figure 1.1 illustrates, as an example, a 2D square Bravais lattice and corresponding reciprocal lattice with the first Brillouin zone and irreducible Brillouin zone.

1.6.2 Three Dimensional Uniform Electron Gas

The uniform electron gas (UEG) is a system consisting of electrons and a neutralizing uniform background. This system is the simplest model for periodic solids and plays an important role in the study of electronic correlations in real systems²⁰. The Hamiltonian of the UEG including N electrons in a cubic box of volume L^3 is as follows,

$$\hat{H} = - \sum_{i=1}^N \frac{1}{2} \nabla_i^2 + \sum_{i=1}^N \sum_{j>i}^N \frac{1}{|\mathbf{r}_i - \mathbf{r}_j|} + \hat{H}_{e-b} + \hat{H}_{b-b}, \quad (1.41)$$

where the last two terms represent the electron-background and background-background interactions, respectively. These two terms fail to converge in the thermodynamic limit $L \rightarrow \infty$; however, this divergence is cancelled by a term from the electron-electron part of the Hamiltonian. In the end, the background interactions are cancelled and the Hamiltonian is dominated by the kinetic energy and Coulomb interaction of electrons. The

second-quantized Hamiltonian of the UEG can be written as

$$\hat{H} = \frac{1}{2} \sum_{\mathbf{k}} k^2 \hat{a}_{\mathbf{k}}^\dagger \hat{a}_{\mathbf{k}} + \sum_{\mathbf{p}, \mathbf{q}} \sum_{\mathbf{k} \neq 0} \frac{4\pi}{k^2} \hat{a}_{\mathbf{p}+\mathbf{k}}^\dagger \hat{a}_{\mathbf{q}-\mathbf{k}}^\dagger \hat{a}_{\mathbf{q}} \hat{a}_{\mathbf{p}}. \quad (1.42)$$

In this form, $\mathbf{k}$, $\mathbf{p}$, and $\mathbf{q}$ refer to reciprocal-space vectors, while $\hat{a}_{\mathbf{k}}^\dagger$ and $\hat{a}_{\mathbf{k}}$ denote the creation and annihilation operators for a single-electron state with a wave vector of $\mathbf{k}$, respectively. The UEG is characterized by its uniform number density n , defined via $r_s = (4\pi n/3)^{-1/3}$, where r_s is called the Wigner-Seitz radius and is defined as the radius of a sphere that encloses, on the average, exactly one electron. It can be proved that the kinetic energy term in the Hamiltonian scales as $1/r_s^2$, while the Coulomb interaction scales as $1/r_s$ ³¹. This scaling property enables us to clearly distinguish two different limiting behaviors of the UEG system. In the high-density regime (small r_s), which is also called the weakly correlated regime, the kinetic energy of the electrons outweighs the potential energy. This creates a fully delocalized system where the UEG behaves as a Fermi liquid. By contrast, in the low-density regime (large r_s), also called the strongly correlated regime, the potential energy dominates over the kinetic energy. As a result, the electrons localize onto lattice points and the system behaves as a classical assembly of charged particles. In this case, the electrons form a crystalline structure that is commonly known as the Wigner crystal³².

1.6.3 Twist Averaging

When one tries to simulate the Hamiltonian of an infinite system using a periodic supercell (a group of unit cells) containing a finite number of electrons, finite-size errors become apparent. The finite size of the supercell imposes restrictions on the minimum momenta of the electrons in reciprocal space. These restrictions ultimately lead to strong fluctuations in calculated properties such as the correlation energy. The resulting finite-size errors are usually substantial and exhibit slow convergence as supercell size increases. A common way to tackle such finite-size errors is to apply the twist averaging procedure^{33–35}. Twist averaging aims to reduce finite-size errors by shifting the $\mathbf{k}$ -point grid in the first Brillouin zone by a small amount $\mathbf{k}_s$, and averaging across all possible offsets. The twist $\mathbf{k}_s$ is defined via $\theta = \mathbf{k}_s \cdot \mathbf{R}$, where $\mathbf{R}$ is a simulation cell lattice vector, and θ is a phase factor originating from the generalized boundary conditions,

$$\Psi(\mathbf{r}_1, \dots, \mathbf{r}_i + \mathbf{R}, \dots, \mathbf{r}_N) = e^{i\theta} \Psi(\mathbf{r}_1, \dots, \mathbf{r}_i, \dots, \mathbf{r}_N), \quad (1.43)$$

which states that electrons are allowed to pick up a phase when they wrap around the periodic boundaries. Consequently, twist averaging makes use of twists as degrees of freedom

of the boundary conditions to provide a smoother extrapolation to the thermodynamic limit. The general form of twist averaging is as follows,

$$A_{\text{TA}} = \frac{1}{\Omega_{\text{BZ}}} \int_{\text{BZ}} A(\mathbf{k}_s) d\mathbf{k}_s, \quad (1.44)$$

where Ω_{BZ} is the volume of the first Brillouin zone, and A is an expectation value that is averaged over $\mathbf{k}_s$ in the Brillouin zone.

Chapter 2

Quantum Chemistry Approaches

This chapter contains a discussion on the most important quantum chemistry approaches to electronic structure calculations, namely Hartree-Fock (HF) theory and the configuration interaction (CI) method^{8,9}. These two approaches make use of the definitions and principles discussed in the previous chapter, and provide a foundation for the next chapter.

2.1 Hartree-Fock Theory

The Hartree-Fock approximation is central to the approximate methods employed to solve the many-electron problem. Additionally, it is usually considered as the first step toward more accurate methods, e.g., post-Hartree-Fock methods and quantum Monte Carlo (QMC) methods. In the Hartree-Fock approximation, a single Slater determinant, which is the simplest antisymmetric wave function, is used for the ground state of an N -electron system.

$$|\Psi_0\rangle = |\chi_1\chi_2\cdots\chi_N\rangle. \quad (2.1)$$

Then the variational principle is used to minimize the ground-state energy,

$$E_0 = \langle\Psi_0|\hat{H}|\Psi_0\rangle, \quad (2.2)$$

where $\hat{H}$ is the electronic Hamiltonian defined in Equation 1.6, and E_0 is the ground-state energy. Minimizing the ground state energy with respect to the choice of spin orbitals χ_i

leads to an eigenvalue equation known as the Hartree-Fock equation, which determines the optimal spin orbitals for the ground-state wave function in Equation 2.1.

$$\hat{f}(i)\chi_a(\mathbf{x}_i) = \varepsilon_a\chi_a(\mathbf{x}_i), \quad (2.3)$$

where $\hat{f}(i)$ is an effective one-body operator, known as the Fock operator, ε_a is the energy of the a th orbital, and i labels the i th electron. The Fock operator consists of two other one-body operators,

$$\hat{f}(i) = \hat{h}(i) + \hat{v}^{\text{HF}}(i), \quad (2.4)$$

where $\hat{h}(i)$, known as the core-Hamiltonian operator, for the i th electron describes the kinetic energy and the potential energy in the field of the nuclei,

$$\hat{h}(i) = -\frac{1}{2}\nabla_i^2 - \sum_{A=1}^M \frac{Z_A}{r_{iA}}. \quad (2.5)$$

$\hat{v}^{\text{HF}}(i)$, known as the Hartree-Fock potential, is also the average potential experienced by the i th electron due to the presence of other electrons, and is defined as

$$\hat{v}^{\text{HF}}(i) = \sum_b^N \hat{J}_b(i) - \hat{K}_b(i), \quad (2.6)$$

where $\hat{J}_b(i)$ and $\hat{K}_b(i)$ are called the Coulomb and exchange operators, respectively, and are defined by

$$\hat{J}_b(i)\chi_a(\mathbf{x}_i) = \left[\int \chi_b^*(\mathbf{x}_j) r_{ij}^{-1} \chi_b(\mathbf{x}_j) d\mathbf{x}_j \right] \chi_a(\mathbf{x}_i), \quad (2.7)$$

$$\hat{K}_b(i)\chi_a(\mathbf{x}_i) = \left[\int \chi_b^*(\mathbf{x}_j) r_{ij}^{-1} \chi_a(\mathbf{x}_j) d\mathbf{x}_j \right] \chi_b(\mathbf{x}_i). \quad (2.8)$$

The Coulomb operator $\hat{J}_b(i)$, as its name suggests, arises from the Coulomb interaction between electrons, and represents the average local one-body Coulomb potential at $\mathbf{x}_i$ from an electron in χ_b . The exchange operator $\hat{K}_b(i)$ arises because of the antisymmetric nature of the determinantal wave function. It is defined by its effect when operating on a spin orbital $\chi_a(\mathbf{x}_i)$ (Equation 2.8), and does not have a simple classical interpretation as the Coulomb term.

Note that the Hartree-Fock potential $\hat{v}^{\text{HF}}(i)$ for the i th electron depends on the spin orbitals of other electrons, as seen in Equation 2.6, meaning that the Fock operator $\hat{f}(i)$ depends on its eigenstates. This makes the Hartree-Fock equation non-linear; consequently, its analytic solution is generally not possible. The self-consistent-field (SCF) method is an approach used to solve the Hartree-Fock equation iteratively. Solving the Hartree-Fock Equation 2.3 gives a set of orthonormal spin orbitals $\{\chi_k\}$ and associated orbital energies $\{\varepsilon_k\}$. The N spin orbitals with the lowest energies are called occupied spin orbitals, and are labeled as $\chi_a, \chi_b, \dots$. The remaining orbitals from the set $\{\chi_k\}$ are called unoccupied or virtual spin orbitals, and are labeled as $\chi_r, \chi_s, \dots$. The Slater determinant formed from the occupied spin orbitals is the Hartree-Fock wave function Ψ_{HF} , and is the best variational solution to the ground state of a system in the single determinantal approximation. Using Equation 2.2, the Hartree-Fock energy would be written as follows,

$$\begin{aligned}
E_{\text{HF}} &= \langle \Psi_{\text{HF}} | \hat{H} | \Psi_{\text{HF}} \rangle \\
&= \left\langle \Psi_{\text{HF}} \left| \sum_{i=1}^N \hat{h}(i) + \sum_{i=1}^N \sum_{j>i}^N r_{ij}^{-1} \right| \Psi_{\text{HF}} \right\rangle \\
&= \sum_a^N \int \chi_a^*(\mathbf{x}_i) \hat{h}(\mathbf{r}_i) \chi_a(\mathbf{x}_i) d\mathbf{x}_i \\
&\quad + \frac{1}{2} \sum_{a,b}^N \int \chi_a^*(\mathbf{x}_i) \chi_b^*(\mathbf{x}_j) r_{ij}^{-1} (1 - \hat{P}_{ij}) \chi_a(\mathbf{x}_i) \chi_b(\mathbf{x}_j) d\mathbf{x}_i d\mathbf{x}_j,
\end{aligned} \tag{2.9}$$

where $\hat{P}_{ij}$ is called the permutation operator, which interchanges the coordinates of the i th and j th electron. In the derivation of the above formula, we have used the Slater-Condon rules³⁶⁻³⁸, which are reviewed in Appendix A. The resulting one-body and two-body integrals can also be written in a short-hand notation known as chemists' notation,

$$E_{\text{HF}} = \sum_a^N [a | \hat{h} | a] + \frac{1}{2} \sum_{a,b}^N [aa | bb] - [ab | ba]. \tag{2.10}$$

All of our formulas so far have involved spin orbitals χ_i rather than spatial orbitals ψ_i . However, in most computational tasks, the spin functions α and β are integrated out in order to reduce the spin-orbital formulations, leaving a spatial-orbital formulation. For an N -electron system containing an even number of electrons, where spatial parts of any spin orbital are restricted to be the same for both α and β spins (i.e., $\{\chi_i\} = \{\psi_i\alpha, \psi_i\beta\}$), namely closed-shell restricted Slater determinant, the Hartree-Fock energy can be written

using one- and two-body integrals over spatial orbitals,

$$E_{\text{HF}} = 2 \sum_a^{N/2} (a|\hat{h}|a) + \sum_{a,b}^{N/2} 2(aa|bb) - (ab|ba) = 2 \sum_a^{N/2} (a|\hat{h}|a) + \sum_{a,b}^{N/2} 2J_{ab} - K_{ab}, \quad (2.11)$$

where round brackets are used to indicate the use of spatial orbitals. J_{ab} and K_{ab} are called the Coulomb integral and exchange integral, respectively, and are given by

$$J_{ab} = (aa|bb) = \int |\psi_a(\mathbf{r}_i)|^2 r_{ij}^{-1} |\psi_b(\mathbf{r}_j)|^2 d\mathbf{r}_i d\mathbf{r}_j, \quad (2.12)$$

$$K_{ab} = (ab|ba) = \int \psi_a^*(\mathbf{r}_i) \psi_b(\mathbf{r}_i) r_{ij}^{-1} \psi_b^*(\mathbf{r}_j) \psi_a(\mathbf{r}_j) d\mathbf{r}_i d\mathbf{r}_j. \quad (2.13)$$

The Coulomb integral J_{ab} is nothing but the classical Coulomb repulsion between electron clouds $|\psi_a(\mathbf{r}_i)|^2$ and $|\psi_b(\mathbf{r}_j)|^2$. The exchange integral K_{ab} is also the result of exchange correlation and does not have any classical interpretation.

The Hartree-Fock equation is solved by considering a finite set of basis functions $\{\phi_\mu(\mathbf{r})|\mu = 1, 2, \dots, K\}$ (see 1.3) in which the orbitals are expanded. The larger the set of basis functions, the greater is the flexibility in the expansion of spin orbitals, and the lower the Hartree-Fock energy. Thus, as the size of the basis set increases, E_{HF} decreases until a limit is reached, called the Hartree-Fock limit. Subsequently, any finite basis set would lead to an energy somewhat larger than the Hartree-Fock limit. As an example, we consider the ground state of a carbon dimer C_2 at equilibrium geometry (C_2 : 2.3481 a_0 ³⁹). We calculate the Hartree-Fock energy of C_2 in Dunning basis sets^{40–42}, cc-pCV n Z ($n=\text{D, T, Q, 5, and 6}$), and then extrapolate it to the complete basis set (CBS) in order to get the Hartree-Fock limit. Using a simple three-parameter fitting function of the cardinal number n with an exponential form⁴³, $E(n) = E(\text{CBS}) + ce^{-\beta n}$, the CBS limit for the HF energies (Figure 2.1) is obtained as -75.4069 Ha. The calculations are done using the MOLPRO code^{44,45}.

2.2 Configuration Interaction

Hartree-Fock theory is a single determinantal approximation, producing a set of $2K$ spin orbitals and yielding the Hartree-Fock energy E_{HF} (Equation 2.11) and Hartree-Fock wave function Ψ_{HF} ,

$$\Psi_{\text{HF}} = |\chi_1 \chi_2 \dots \chi_a \chi_b \dots \chi_N\rangle. \quad (2.14)$$

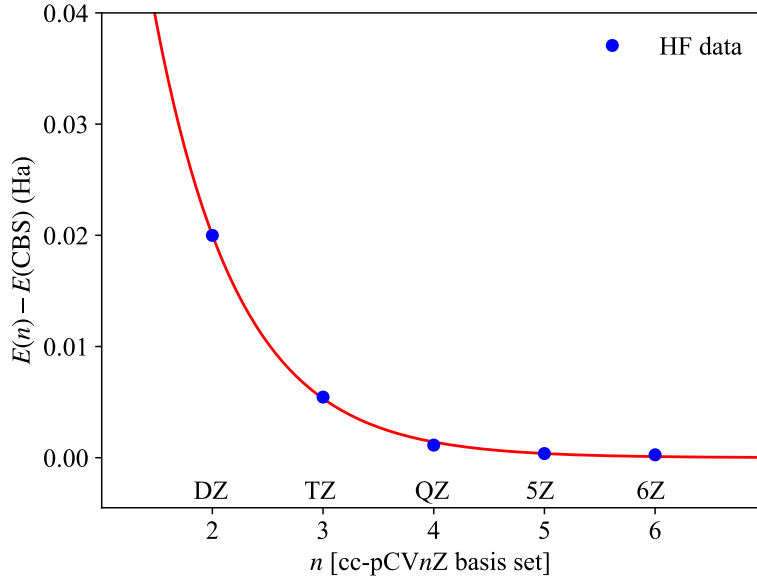


Figure 2.1: CBS extrapolation of the Hartree-Fock energy of a carbon dimer C_2 at equilibrium geometry.

In general, one can construct different individual determinants from N electrons and $2K > N$ spin orbitals. The number of different determinants would be equal to the combination of N electrons from a set of $2K$ spin orbitals,

$$\binom{2K}{N} = \frac{(2K)!}{N!(2K - N)!}, \quad (2.15)$$

where the Hartree-Fock determinant is just one of those. The other determinants can also be classified with respect to the Hartree-Fock determinant. For example, a singly-excited determinant $|\Psi_a^r\rangle$ (Figure 2.2) would be the one in which an electron in occupied spin orbital χ_a has been promoted to a virtual spin orbital χ_r . All $\binom{2K}{N}$ determinants can then be categorized as either the Hartree-Fock determinant or a singly-, doubly-, triply-, ... excited determinant. Although excited determinants are rarely considered as valid excited states wave function of the system as they are far less accurate than the Hartree-Fock determinant, they constitute an N -electron basis for describing the exact states of the system.

Besides, there are some methods, known as post-Hartree-Fock methods, which attempt to use more than one Slater determinant in the description of the wave function. Espe-

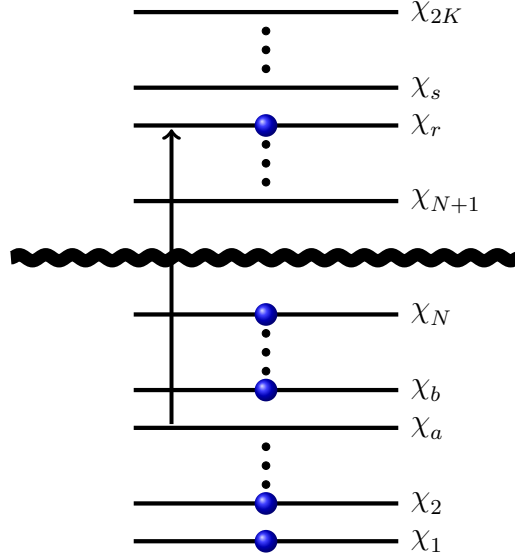


Figure 2.2: A singly-excited determinant $|\Psi_a^r\rangle$.

cially for electrons with antiparallel spins, these methods are typically more accurate than Hartree-Fock theory, but they are more computationally expensive as well. Configuration interaction is one of the most popular post-Hartree-Fock methods, which is based on the fact that the exact wave function for the ground and excited states of an N -electron system can be written as a linear combination of all possible Slater determinants constructed from a complete set of spin orbitals,

$$|\Phi\rangle = c_0|\Psi_0\rangle + \sum_{ra} c_a^r |\Psi_a^r\rangle + \sum_{\substack{a<b \\ r<s}} c_{ab}^{rs} |\Psi_{ab}^{rs}\rangle + \sum_{\substack{a<b<c \\ r<s<t}} c_{abc}^{rst} |\Psi_{abc}^{rst}\rangle + \dots, \quad (2.16)$$

where on the right-hand side, the summations are over all unique singly-, doubly-, triply-, ... excited determinants, respectively. Given a complete set of spin orbitals $\{\chi_i\}$, the infinite set of N -electron determinants $\{|\Psi_i\rangle\}$ would form a complete set for the expansion of any N -electron wave function. Then, one can obtain the exact ground and excited states energies by diagonalizing the Hamiltonian matrix formed from the complete set $\{|\Psi_i\rangle\}$. The lowest eigenvalue of the Hamiltonian matrix $\mathcal{E}_0$ is the exact non-relativistic ground-state energy of the system within the Born-Oppenheimer approximation. The difference between this energy and the Hartree-Fock limit E_0 is called the correlation energy, as the motion

of electrons with opposite spins is not correlated within the Hartree-Fock theory.

$$E_{\text{corr}} = \mathcal{E}_0 - E_0. \quad (2.17)$$

In practice, one cannot perform the above procedure as it is impossible to handle an infinite basis set. However, we can work with a finite set of spin orbitals instead as the one-electron basis set. The resulting N -electron basis set then will not be complete. Nevertheless, we can always have exact solutions within the subspace spanned by the finite set of N -electron basis functions. This procedure is called full configuration interaction (FCI). The number of determinants in the FCI expansion even in small systems and small basis sets is so large, thus in practice one must truncate the expansion and use a small fraction of $\binom{2K}{N}$ determinants.

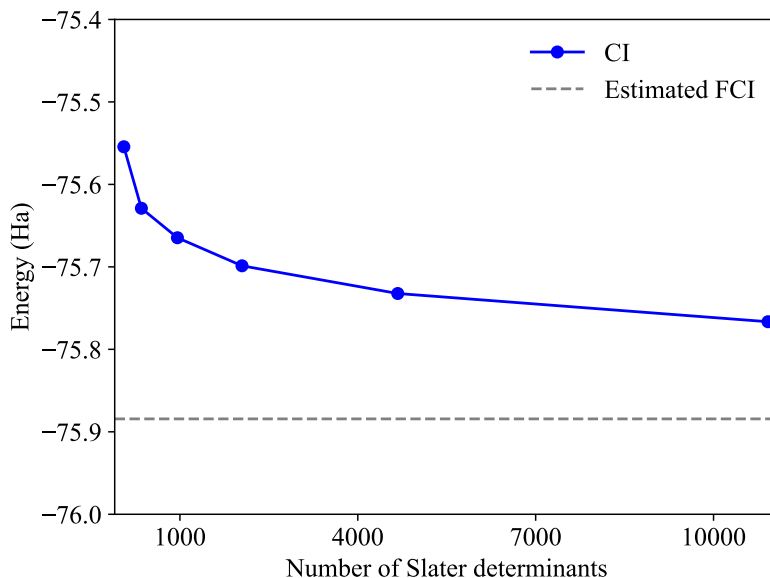


Figure 2.3: Variation of CI energy of a carbon dimer C_2 at equilibrium geometry using the cc-pCVTZ basis with the number of determinants, including the estimated FCI energy from an FCIQMC calculation.

Figure 2.3 shows the CI energies of a carbon dimer C_2 at equilibrium geometry in cc-pCVTZ basis set as a function of the number of determinants in the CI expansion. As it can be seen, the CI energy decreases with increasing number of determinants, and it will eventually converge to the FCI energy when all the $\binom{2K}{N}$ determinants are included

in the expansion. To estimate the FCI energy, we use the very accurate and very efficient FCIQMC method^{12,13}, as the direct FCI calculation would require a vast amount of storage.

2.2.1 Heat-Bath Configuration Interaction

The number of determinants in an FCI expansion scales combinatorially with the number of electrons and orbitals, however many of these determinants do not contribute effectively to retrieve the correlation energy, and hence are considered as deadwood configurations⁴⁶. Taking advantage of this fact to select only the important determinants introduces a wide range of approaches, namely the selected configuration interaction (SCI) methods^{22,47–56}. Selected configuration interaction plus perturbation theory (SCI+PT) algorithms are a subclass of these methods which improve upon the variational energy by employing the multi-reference Epstein-Nesbet perturbation theory^{57,58}. CIPSI (configuration interaction by perturbation with multi-configurational zeroth-order wave function selected by iterative process)⁴⁷ is known as the first such techniques.

Heat-bath configuration interaction is a SCI+PT method which employs an algorithm that greatly improves the efficiency of both the variational and perturbative steps⁵⁵. This method consists of two stages, namely a variational stage and a perturbative stage. During the first stage, a variational wave function is constructed as a linear combination of a set of determinants which are chosen iteratively. Then in the second stage, a second-order energy correction is applied using the multi-reference Epstein-Nesbet perturbation theory. To begin with, one needs to define the variational space $\mathcal{V}$, where the Hamiltonian is diagonalized, returning the ground-state wave function and energy. The space of connections $\mathcal{C}$ is also defined as a space of determinants which are connected to $\mathcal{V}$ but are not in $\mathcal{V}$. The algorithm starts with a variational space $\mathcal{V}$ consisting of an initial set of determinants, or just the Hartree-Fock determinant. It then iterates as follows,

1. Calculate the lowest eigenvalue E_0 with eigenvector $|\Psi_0\rangle = \sum_i c_i |D_i\rangle$ of the Hamiltonian in the variational space $\mathcal{V}$.
2. Find all determinants $\{D_a\}$ in the space of connections $\mathcal{C}$ for which $|H_{ai}c_i| > \epsilon_1$ for at least one determinant D_i in the variational space $\mathcal{V}$.
3. Add those determinants to the variational space $\mathcal{V}$.
4. Repeat steps 1 through 4 until an iteration is reached in which the number of new determinants is less than a threshold, for example, 1% of the number of determinants already present in $\mathcal{V}$.

The above iteration requires the parameter ϵ_1 as a threshold to add the new determinants to the variational space. The Hamiltonian matrix element $\langle D_a | \hat{H} | D_i \rangle$ is shortened as H_{ai} , and the indices $i, j, \dots$ and $a, b, \dots$ are used for determinants in the variational space and space of connections, respectively. Then in the perturbative stage, a second-order correction to the variational energy is computed as follows,

$$\Delta E_2 \approx \sum_a \frac{(\sum_i^{(\epsilon_2)} H_{ai} c_i)^2}{E_0 - H_{aa}}, \quad (2.18)$$

where ϵ_2 is called the perturbative threshold, and $\sum_i^{(\epsilon_2)}$ denotes a summation over those variational determinants for which $|H_{ai} c_i| > \epsilon_2$. The HCI method makes use of the parameter ϵ_2 to include only those terms in the above summation that contribute substantially. The perturbative threshold ϵ_2 is usually set much smaller than the variational threshold ϵ_1 as discarding small amplitude determinants can lead to large errors in the calculation of dynamical correlation. Also, the smaller the perturbative threshold ϵ_2 , the more is the number of determinants in the perturbative space, which can lead to a severe memory bottleneck. To avoid this issue, one can use a stochastic or semistochastic version of the perturbative correction²². In the stochastic version, the expectation value of the second-order energy correction (Equation 2.18) is calculated stochastically. This is done by taking N_s samples, each consisting of N_d determinants $\{D_i\}$ sampled from the variational space $\mathcal{V}$ with the following probability,

$$p_i = \frac{|c_i|}{\sum_i |c_i|}. \quad (2.19)$$

Each sample contains N_d^{dist} distinct determinants D_i , each with a repetition number w_i , such that

$$\sum_i^{N_d^{\text{dist}}} w_i = N_d. \quad (2.20)$$

Using the above definitions and applying some simplifications, the second-order correction can be estimated as follows,

$$\begin{aligned} \Delta E_2 = & \frac{1}{N_d(N_d - 1)} \left\langle \sum_a \frac{1}{E_0 - H_{aa}} \left[\left(\sum_i^{N_d^{\text{dist}}(\epsilon_2)} \frac{w_i c_i H_{ai}}{p_i} \right)^2 \right. \right. \\ & \left. \left. + \sum_i^{N_d^{\text{dist}}(\epsilon_2)} \left(\frac{w_i(N_d - 1)}{p_i} - \frac{w_i^2}{p_i^2} \right) c_i^2 H_{ai}^2 \right] \right\rangle. \end{aligned} \quad (2.21)$$

Reducing the stochastic error to a desired value requires large number of samples N_s to be evaluated, making the stochastic method rather expensive for relatively large variational spaces and large basis sets. The semistochastic method reduces the computational cost by splitting the perturbative correction into two steps. During the first step, a second-order energy correction $\Delta E_2^D[\epsilon_2^d]$ is calculated deterministically using a relatively large threshold ϵ_2^d . The error originating from the use of large ϵ_2^d is then corrected stochastically in the second step. This is achieved by calculating the difference between the second-order corrections obtained from a small threshold ϵ_2 and a large threshold ϵ_2^d as follows,

$$\Delta E_2[\epsilon_2] = (\Delta E_2^S[\epsilon_2] - \Delta E_2^S[\epsilon_2^d]) + \Delta E_2^D[\epsilon_2^d], \quad (2.22)$$

where $\Delta E_2^S[\epsilon_2]$ and $\Delta E_2^S[\epsilon_2^d]$ are the second-order energy corrections calculated stochastically using ϵ_2 and ϵ_2^d , respectively.

Chapter 3

Quantum Monte Carlo Approaches

This chapter focuses on solving the Schrödinger equation using the Monte Carlo method⁵⁹, namely quantum Monte Carlo (QMC). The term “quantum Monte Carlo” covers several different techniques based on random sampling, such as variational Monte Carlo (VMC)^{10,14}, diffusion Monte Carlo (DMC)^{11,14}, and full configuration interaction quantum Monte Carlo (FCIQMC)^{12,13}. QMC calculations are expensive, which limits their applicability at present; however, they are the most accurate methods known so far for computing the energies of large assemblies of interacting quantum particles. The high accuracy of QMC methods is essential in problems in which a precise description of the underlying physics is required.

3.1 The Metropolis Algorithm

The Metropolis algorithm⁶⁰ is a Markov chainⁱ Monte Carlo (MCMC) method to obtain a sequence of random samples from a probability distribution for which direct sampling is difficult. The algorithm generates a sequence of sampling points $\mathbf{R}_i$ by moving a single walker as follows,

1. Start the walker at a random position $\mathbf{R}$.

ⁱMarkov chain is a stochastic model describing a sequence of random variables with the important feature that each element has memory of its immediate predecessor. A simple example for a Markov chain is a random walk.

2. Make a trial move to a new position $\mathbf{R}'$ drawn from the probability density $T(\mathbf{R}' \leftarrow \mathbf{R})$.
3. Accept the trial move to $\mathbf{R}'$ with probability $A(\mathbf{R}' \leftarrow \mathbf{R}) = \text{Min}\left[1, \frac{T(\mathbf{R} \leftarrow \mathbf{R}')\mathcal{P}(\mathbf{R}')}{T(\mathbf{R}' \leftarrow \mathbf{R})\mathcal{P}(\mathbf{R})}\right]$.
4. Return to step 2.

If the trial move at step 3 is accepted, the point $\mathbf{R}'$ becomes the next position of the walker; otherwise, the walker remains at $\mathbf{R}$. It can be readily verified that this algorithm satisfies the detailed balance condition,

$$\mathcal{P}(\mathbf{R})T(\mathbf{R}' \leftarrow \mathbf{R})A(\mathbf{R}' \leftarrow \mathbf{R}) = \mathcal{P}(\mathbf{R}')T(\mathbf{R} \leftarrow \mathbf{R}')A(\mathbf{R} \leftarrow \mathbf{R}'). \quad (3.1)$$

Hence $\mathcal{P}(\mathbf{R})$ is the equilibrium distribution of this Markov process. As long as the Markov process is ergodic (i.e., it is possible to reach any point in configuration space in a finite number of moves), it can be shown that the process will converge to this equilibrium distribution.

3.2 Variational Monte Carlo

The variational Monte Carlo (VMC) method^{10,14} is based on a combination of the variational principle and the Monte Carlo evaluation of integrals. The energy is calculated as the expectation value of the Hamiltonian with a real, properⁱⁱ many-body trial wave function. In general, there are no restrictions on the form of the trial wave function. The variational principle for the trial wave function Ψ_T states as

$$E_V = \frac{\int \Psi_T^*(\mathbf{R}) \hat{H} \Psi_T(\mathbf{R}) d\mathbf{R}}{\int \Psi_T^2(\mathbf{R}) d\mathbf{R}} \geq E_0, \quad (3.2)$$

where E_0 is the exact ground-state energy, $\hat{H}$ is the many-body Hamiltonian, and $\mathbf{R}$ is a $3N$ -dimensional vector of particle coordinates. To integrate E_V using the Monte Carlo method, an importance sampling transformation of which can be written as

$$E_V = \int \mathcal{P}(\mathbf{R}) E_L(\mathbf{R}) d\mathbf{R}, \quad (3.3)$$

ⁱⁱThe trial wave function Ψ_T must have the correct symmetry under particle exchange; Both Ψ_T and $\nabla \Psi_T$ must be continuous wherever the potential is finite; The integrals $\int \Psi_T^* \Psi_T d\mathbf{R}$ and $\int \Psi_T^* \hat{H} \Psi_T d\mathbf{R}$ must exist, and to keep the variance of the energy finite it is desirable that $\int \Psi_T^* \hat{H}^2 \Psi_T d\mathbf{R}$ exists.⁶¹

where $\mathcal{P}(\mathbf{R})$ is an importance function, which is defined as

$$\mathcal{P}(\mathbf{R}) = \frac{\Psi_T^2(\mathbf{R})}{\int \Psi_T^2(\mathbf{R}') d\mathbf{R}'}, \quad (3.4)$$

and $E_L(\mathbf{R})$ is called the local energy,

$$E_L(\mathbf{R}) = \frac{\hat{H}\Psi_T(\mathbf{R})}{\Psi_T(\mathbf{R})}, \quad (3.5)$$

which can be evaluated at any $\mathbf{R}$. Since the importance function $\mathcal{P}(\mathbf{R})$ satisfies the following equations,

$$\mathcal{P}(\mathbf{R}) \geq 0 \quad (3.6)$$

$$\int \mathcal{P}(\mathbf{R}) d\mathbf{R} = 1, \quad (3.7)$$

it can be regarded as a probability distribution. Therefore, the value of E_V may be obtained by drawing an infinite set of random vectors from the probability distribution $\mathcal{P}(\mathbf{R})$ and computing the sample average as follows,

$$E_V = \lim_{M \rightarrow \infty} \left[\frac{1}{M} \sum_{i=1}^M E_L(\mathbf{R}_i) \right], \quad (3.8)$$

where M denotes the number of sample vectors, or the number of configurations (also called “walkers”). A Monte Carlo estimate of E_V can also be obtained by averaging over a large but finite number of vectors drawn from $\mathcal{P}(\mathbf{R})$,

$$E_V \approx \frac{1}{M} \sum_{i=1}^M E_L(\mathbf{R}_i), \quad (3.9)$$

where M configurations have been generated after equilibration. The Metropolis algorithm (see 3.1) is used to sample the set of finite vectors $\{\mathbf{R}_i | i = 1, 2, \dots, M\}$ from the probability distribution $\mathcal{P}(\mathbf{R})$. The trial moves are usually sampled from a Gaussian centered on the current position of the walker, and the variance of the Gaussian is chosen such that the average acceptance ratio is roughly 50% or such that the diffusion constant of the random walk is maximized⁶¹. Figure 3.1 illustrates the flow chart of the VMC algorithm.

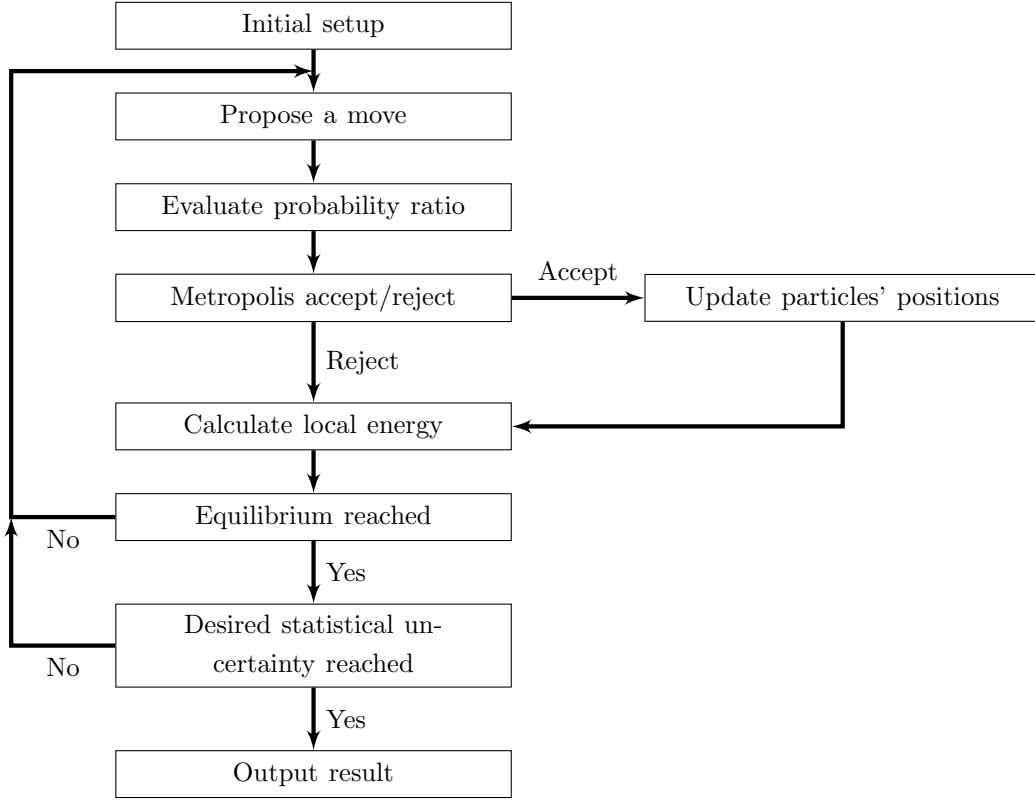


Figure 3.1: Flow chart of the VMC algorithm.

Moreover, the central limit theorem tells us that for large enough M , the variance of E_V is $\sigma_{E_L}^2/M$, which is estimated as

$$\frac{\sigma_{E_L}^2}{M} \approx \frac{1}{M(M-1)} \sum_{i=1}^M \left[E_L(\mathbf{R}_i) - \frac{1}{M} \sum_{j=1}^M E_L(\mathbf{R}_j) \right]^2. \quad (3.10)$$

Hence, $\pm\sigma_{E_L}/\sqrt{M}$ is as an estimate of the size of the error bar on E_V . We use VMC mostly for optimizing parameters in trial wave functions, and our main part of calculations are done with more sophisticated diffusion Monte Carlo (DMC) method, which is discussed in the next section.

3.3 Diffusion Monte Carlo

Diffusion Monte Carlo (DMC) ^{11,14} is a stochastic projector method which provides a practical way of evolving the wave function of a quantum system in imaginary time and obtaining, ultimately, the ground- and excited-states energy ⁱⁱⁱ of the system.

3.3.1 The Imaginary-time Schrödinger Equation

To begin with, consider the time-dependent Schrödinger equation,

$$\hat{H}\Phi(\mathbf{R}, t) = i\frac{\partial\Phi(\mathbf{R}, t)}{\partial t}, \quad (3.11)$$

where $\Phi(\mathbf{R}, t)$ is the time-dependent wave function, and $\hat{H}$ is the electronic Hamiltonian (see 1.6). The formal solution to this equation can be written using the eigenfunctions of $\hat{H}$,

$$\Phi(\mathbf{R}, t) = \sum_{n=0}^{\infty} c_n \Phi_n(\mathbf{R}) e^{-iE_n t}, \quad (3.12)$$

where $\Phi_n(\mathbf{R})$ and E_n are the eigenfunctions and the eigenvalues of the time-independent Schrödinger equation 1.3, respectively. The eigenfunctions $\Phi_n(\mathbf{R})$ are assumed to be orthonormal and real,

$$\int_{-\infty}^{\infty} \Phi_n(\mathbf{R}) \Phi_m(\mathbf{R}) d\mathbf{R} = \delta_{nm}. \quad (3.13)$$

Using the orthonormality condition, the expansion coefficients in Equation 3.12 can be written as

$$c_n = \int_{-\infty}^{\infty} \Phi_n(\mathbf{R}) \Phi(\mathbf{R}, 0) d\mathbf{R} \quad n = 0, 1, 2, \dots \quad (3.14)$$

The coefficients c_n then describe the overlap of the initial state $\Phi(\mathbf{R}, 0)$ with the eigenfunctions $\Phi_n(\mathbf{R})$. Now, a methodological shift of energy scale can be performed,

$$(\hat{H} - E_T)\Phi(\mathbf{R}, t) = i\frac{\partial\Phi(\mathbf{R}, t)}{\partial t}, \quad (3.15)$$

ⁱⁱⁱThe fixed-node DMC (discussed later, see 3.3.2) is always variational for the ground state, ensuring the convergence to the lowest energy state having an energy always greater than or equal to the exact ground-state energy. By contrast, the existence of a variational principle for excited-states cannot in general be guaranteed, and it depends on the symmetry of the trial wave function ⁶².

which leads to the solution

$$\Phi(\mathbf{R}, t) = \sum_{n=0}^{\infty} c_n \Phi_n(\mathbf{R}) e^{-i(E_n - E_T)t}. \quad (3.16)$$

A transformation from real time to imaginary time (known as Wick rotation) is then performed. Introducing a new variable $\tau = it$ as the imaginary time, the Schrödinger equation is rewritten as

$$(\hat{H} - E_T)\Phi(\mathbf{R}, \tau) = -\frac{\partial \Phi(\mathbf{R}, \tau)}{\partial \tau}. \quad (3.17)$$

This is called the imaginary-time Schrödinger equation, which provides the fundamental starting point of all projector quantum Monte Carlo methods. Equation 3.16 is then rewritten as

$$\Phi(\mathbf{R}, \tau) = \sum_{n=0}^{\infty} c_n \Phi_n(\mathbf{R}) e^{-(E_n - E_T)\tau}. \quad (3.18)$$

Now let us investigate the asymptotic behavior of $\Phi(\mathbf{R}, \tau)$ for $\tau \rightarrow \infty$,

1. if $E_T > E_0$, $\lim_{\tau \rightarrow \infty} \Phi(\mathbf{R}, \tau) = \infty$, the wave function diverges exponentially fast.
2. if $E_T < E_0$, $\lim_{\tau \rightarrow \infty} \Phi(\mathbf{R}, \tau) = 0$, the wave function vanishes exponentially fast.
3. if $E_T = E_0$, $\lim_{\tau \rightarrow \infty} \Phi(\mathbf{R}, \tau) = c_0 \Phi_0(\mathbf{R})$, the wave function converges to the ground state exponentially fast.

In the case of $E_T = E_0$, $\Phi(\mathbf{R}, \tau)$ converges to the ground-state wave function $\Phi_0(\mathbf{R})$ regardless of the choice of the initial wave function $\Phi(\mathbf{R}, 0)$, as long as there is a numerically significant overlap between $\Phi(\mathbf{R}, 0)$ and $\Phi_0(\mathbf{R})$; in other words, as long as c_0 is not too small. The aspiration of all projector methods is to realize this long-time limit by performing a stochastic integration of Equation 3.17. The Hamiltonian in Equation 3.17 is the sum of kinetic and potential terms: $\hat{H} = -\sum_{i=1}^N (1/2) \nabla_i^2 + V(\mathbf{R})$. If we neglect the potential term, Equation 3.17 reduces to a diffusion equation. If, on the other hand, we neglect the kinetic term, 3.17 reduces to a rate equation. The Equation 3.17 may also be written in the integral form,

$$\Phi(\mathbf{R}, t + \tau) = \int G(\mathbf{R} \leftarrow \mathbf{R}', \tau) \Phi(\mathbf{R}', t) d\mathbf{R}', \quad (3.19)$$

where t is the imaginary time variable and $G(\mathbf{R} \leftarrow \mathbf{R}', \tau)$ is a Green's function, that is a solution of Equation 3.17 with initial condition $G(\mathbf{R} \leftarrow \mathbf{R}', 0) = \delta(\mathbf{R} - \mathbf{R}')$. The exact Green's function can be sampled using the Green's function Monte Carlo method (GFMC)^{63,64}. The GFMC method is computationally expensive and one can instead approximate the Green's function using the short-time approximation,

$$G(\mathbf{R} \leftarrow \mathbf{R}', \tau) \approx G_d(\mathbf{R} \leftarrow \mathbf{R}', \tau) G_b(\mathbf{R} \leftarrow \mathbf{R}', \tau), \quad (3.20)$$

where

$$G_d(\mathbf{R} \leftarrow \mathbf{R}', \tau) = (2\pi\tau)^{-3N/2} \exp \left[-\frac{(\mathbf{R} - \mathbf{R}')^2}{2\tau} \right], \quad (3.21)$$

is the diffusion Green's function and

$$G_b(\mathbf{R} \leftarrow \mathbf{R}', \tau) = \exp \left[-\frac{\tau}{2} (V(\mathbf{R}) + V(\mathbf{R}') - 2E_T) \right], \quad (3.22)$$

is the branching term. Now using the short-time approximation, one can simulate the imaginary-time evolution by taking a population of configurations $\{\mathbf{R}_i\}$ and subjecting them to random moves drawn from $G_d(\mathbf{R} \leftarrow \mathbf{R}', \tau)$ to simulate the diffusion process. Then, one can use the branching or birth-death algorithm to simulate the rate process, where $G_b(\mathbf{R} \leftarrow \mathbf{R}', \tau)$ acts as a time-dependent renormalization factor, which determines the number of walkers that survive to the next random move.

In DMC, the integration is done in the particle coordinate space, which is a symmetrized space, and then admits solutions with bosonic character (i.e., nodeless and symmetric wave functions). This issue is known as the “fermion sign problem” or “boson catastrophe”, which stems from the antisymmetry property of many-electron wave functions to electron exchange, that leads to wave functions having both positive and negative amplitudes. However, this problem can be evaded to some extent by using the fixed-node approximation^{65–67}.

3.3.2 The Fixed-node Approximation

The basic idea to stochastically integrate Equation 3.17 in DMC is to consider the wave function itself to be a probability density (of walkers distributed in the electron coordinate space). This implies that the wave function must be a positive definite function. For a single fermion, the ground-state wave function has no nodes and one can always fulfill the requirement of overlap between $\Phi(\mathbf{R}, 0)$ and $\Phi_0(\mathbf{R})$ by choosing a positive definite $\Phi(\mathbf{R}, 0)$

centered in a region of space where $\Phi_0(\mathbf{R})$ is sufficiently large. For two or more identical fermions, the wave function must have both positive and negative regions, as should an excited state of any system. Hence, the wave function cannot in general be considered as a probability density.

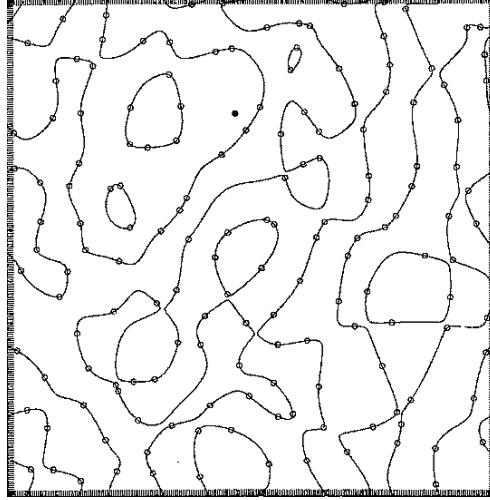


Figure 3.2: A 2-dimensional slice through the 321-dimensional nodal surface of a 2-dimensional Fermi gas containing 161 (polarized) fermions (reproduced with permission from Springer Nature. © Plenum Publishing Corporation 1991⁴). All 161 particle positions were sampled using VMC. The filled circle indicates the original position of the first particle. The other 160 particles are fixed at positions indicated by the open circles. The nodes of the wave function as a function of the position of the first particle are plotted.

The fixed-node approximation circumvents this difficulty by using an approximate nodal surface. In a 3D system containing N electrons, the trial nodal surface is a $(3N-1)$ -dimensional surface on which the wave function is zero and across which it changes sign. Figure 3.2 illustrates a 2D cross section of the nodal surface of a 2D Fermi gas containing 161 (polarized) fermions, where 160 fermions are fixed at a typical (random) position⁴. The fixed-node approximation is equivalent to placing an infinite repulsive potential barrier on the nodal surface of the trial wave function, which is sufficiently strong to force the wave function to be zero on the nodal surface. Consequently, the Schrödinger equation is solved within each pocket enclosed by the nodal surface. The fixed-node DMC method then projects out, within each pocket, the lowest-energy nodeless wave function satisfying zero boundary conditions on the enclosing nodal surface. The implementation of the fixed-node approximation algorithm is also straightforward. The only new ingredient to

the DMC algorithm is that, after every DMC (walker) move the sign of the trial wave function is checked and the walker is deleted if it has crossed the trial nodal surface. The infinite repulsive potential barrier has no effect if the trial nodal surface is placed correctly, otherwise the energy is always raised. Therefore, the fixed-node DMC may be regarded as a variational method that gives exact results if the trial nodal surface is exact.

During the fixed-node DMC simulation, the DMC walkers will eventually concentrate in the nodal pockets of the lowest energy class.^{iv} Within any such pocket v_α , there is a unique ground-state eigenfunction $\Phi_0^\alpha(\mathbf{R})$ with eigenvalue E_0^α , which satisfies the following equations,

$$\hat{H}\Phi_0^\alpha(\mathbf{R}) = E_0^\alpha\Phi_0^\alpha(\mathbf{R}) \quad \mathbf{R} \in v_\alpha \quad (3.23)$$

$$\Phi_0^\alpha(\mathbf{R}) = 0 \quad \mathbf{R} \notin v_\alpha. \quad (3.24)$$

Moreover, it can be proved that, each of the eigenvalues E_0^α is an upper bound to the exact ground-state energy. This is called the fixed-node variational principle^{67,68}.

3.3.3 Importance Sampling

The fixed-node DMC algorithm outlined in the previous section is extremely inefficient and gives unsatisfactory results. The inefficiency problem can be tackled by carrying out an importance-sampling transformation using a trial wave function $\Psi_T(\mathbf{R})$ ^{61,67,69}. To this end, one can multiply Equation 3.17 by $\Psi_T(\mathbf{R})$ and then introduce the following function,

$$f(\mathbf{R}, t) = \Psi_T(\mathbf{R})\Phi(\mathbf{R}, t), \quad (3.25)$$

which always has the same sign if and only if the nodal surface of $\Phi(\mathbf{R}, t)$ matches exactly that of $\Psi_T(\mathbf{R})$. Rearranging terms in Equation 3.17 yields

$$-\frac{\partial f}{\partial t} = -\frac{1}{2}\nabla_{\mathbf{R}}^2 f + \nabla_{\mathbf{R}} \cdot [\mathbf{v}_D f] + [E_L - E_T]f, \quad (3.26)$$

where $\mathbf{v}_D$ is the 3-dimensional drift velocity defined as

$$\mathbf{v}_D(\mathbf{R}) = \nabla_{\mathbf{R}} \ln |\Psi_T(\mathbf{R})| = \Psi_T^{-1}(\mathbf{R}) \nabla_{\mathbf{R}} \Psi_T(\mathbf{R}), \quad (3.27)$$

^{iv} A pocket class refers to a group of pockets related together by the permutation symmetry. Furthermore, the tiling theorem states that all the ground-state (lowest-energy) nodal pockets belong to the same class^{4,61}.

and $E_L(\mathbf{R}) = \Psi_T^{-1}(\mathbf{R}) \hat{H} \Psi_T(\mathbf{R})$ is the local energy as defined in VMC. The three terms on the right-hand side of Equation 3.26 correspond to diffusion, drift, and branching processes, respectively. The corresponding integral equation of 3.26 is written as

$$f(\mathbf{R}, t + \tau) = \int G(\mathbf{R} \leftarrow \mathbf{R}', \tau) f(\mathbf{R}', t) d\mathbf{R}', \quad (3.28)$$

and the short-time approximation to $G(\mathbf{R} \leftarrow \mathbf{R}', \tau)$, as in 3.20, is written as $G(\mathbf{R} \leftarrow \mathbf{R}', \tau) \approx G_d(\mathbf{R} \leftarrow \mathbf{R}', \tau) G_b(\mathbf{R} \leftarrow \mathbf{R}', \tau)$, where

$$G_d(\mathbf{R} \leftarrow \mathbf{R}', \tau) = (2\pi\tau)^{-3N/2} \exp \left[- \frac{(\mathbf{R} - \mathbf{R}' - \tau \mathbf{v}_D(\mathbf{R}'))^2}{2\tau} \right], \quad (3.29)$$

$$G_b(\mathbf{R} \leftarrow \mathbf{R}', \tau) = \exp \left[- \frac{\tau}{2} (E_L(\mathbf{R}) + E_L(\mathbf{R}') - 2E_T) \right]. \quad (3.30)$$

The importance sampling transformation has several consequences. First, the density of walkers increases in regions where $|\Psi_T|$ is large and vice versa. This is due to the drift velocity $\mathbf{v}_D$, which carries the walkers along in the direction of increasing $|\Psi_T|$, causing more important parts of the wave function to be sampled. Second, the renormalization factor 3.30 now contains the local energy, which is a much smoother function than the potential energy; thus, the population of walkers is better controlled. Moreover, the importance sampling is a great choice to satisfy the fixed-node constraint. Whenever a walker approaches the nodal surface of the trial wave function, the drift velocity suddenly grows and carries it away.

The short-time approximation makes the DMC results dependent on the time step. To ensure the accuracy of the short-time approximation, one should use time steps τ as small as possible. It is also common to linearly extrapolate the energy to the zero time-step limit^{14,70}. Figure 3.3 shows the linear extrapolation for a carbon dimer at equilibrium geometry, in which fitting is carried out via a custom regression tool called POLYFIT⁷¹. This is a toolbox which uses Monte Carlo resampling to perform and manipulate polynomial fits to data with and without statistical uncertainties. Throughout this thesis, POLYFIT is widely used for performing fits and extrapolations. Besides, all VMC and DMC calculations are done using the CASINO code^{14,72} in this thesis.

The approximate Green's functions in Equations 3.29 and 3.30 are reasonable approximations in regions where Ψ_T is smooth and non-zero. This is not the case however in regions close to the nodal surface, where quantities such as the drift velocity and the local energy may diverge⁷³. A good remedy is to consider a rejection step for the random moves

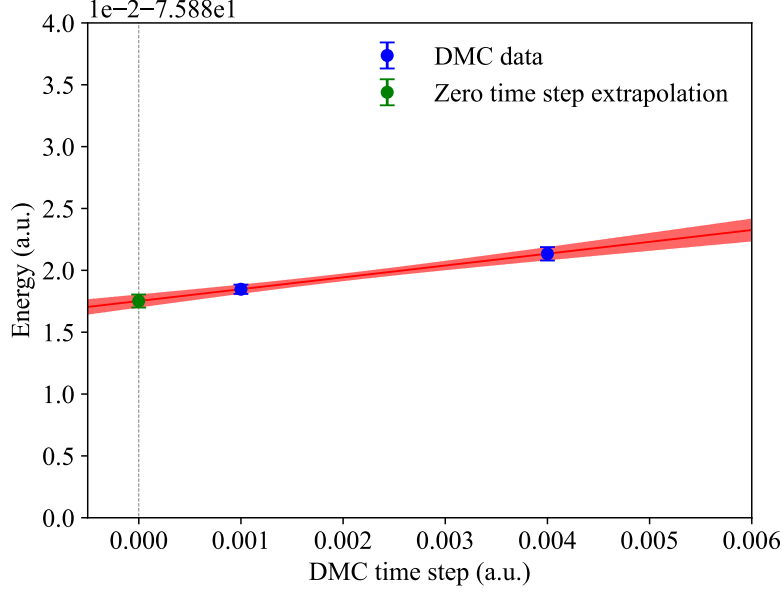


Figure 3.3: Extrapolation of the DMC energy with respect to the time step for a carbon dimer C_2 at equilibrium geometry using a Slater-Jastrow trial wave function. The curve is a linear fit of the data. The width of the line reflects the statistical uncertainty of the fit values at each point.

drawn from $G_d(\mathbf{R} \leftarrow \mathbf{R}', \tau)$. Given that, a trial move from $\mathbf{R}'$ to $\mathbf{R}$ is accepted with probability

$$\begin{aligned}
 A(\mathbf{R} \leftarrow \mathbf{R}') &= \min \left[1, \frac{G_d(\mathbf{R}' \leftarrow \mathbf{R}, \tau) G_b(\mathbf{R}' \leftarrow \mathbf{R}, \tau) \Psi_T^2(\mathbf{R})}{G_d(\mathbf{R} \leftarrow \mathbf{R}', \tau) G_b(\mathbf{R} \leftarrow \mathbf{R}', \tau) \Psi_T^2(\mathbf{R}')} \right] \\
 &= \min \left[1, \frac{G_d(\mathbf{R}' \leftarrow \mathbf{R}, \tau) \Psi_T^2(\mathbf{R})}{G_d(\mathbf{R} \leftarrow \mathbf{R}', \tau) \Psi_T^2(\mathbf{R}')} \right].
 \end{aligned} \tag{3.31}$$

Hence, the fixed-node DMC algorithm reduces to the VMC algorithm with trial moves sampled from $G_d(\mathbf{R} \leftarrow \mathbf{R}', \tau)$. The acceptance probability depends on the time step, and tends to unity as the time step goes to zero. If the time step is large, too many moves are rejected from the beginning and it will bias the sampling of $f(\mathbf{R}, t)$. Then, one should choose the time step such that the acceptance ratio is $\geq 99\%$. The importance-sampled

DMC algorithm finally generates configurations asymptotically distributed according to

$$f(\mathbf{R}, \infty) = \Psi_T(\mathbf{R})\Phi_0(\mathbf{R}), \quad (3.32)$$

where $\Phi_0(\mathbf{R})$ is the ground-state eigenfunction subjected to the fixed-node boundary condition, and is defined as

$$\Phi_0(\mathbf{R}) = \sum_{\alpha} c_{\alpha} \Phi_0^{\alpha}(\mathbf{R}), \quad (3.33)$$

where c_{α} is a non-negative constant proportional to the initial population of the nodal packet v_{α} (see 3.3.2). The fixed-node DMC energy can then be evaluated as

$$E_D = \frac{\langle \Phi_0 | \hat{H} | \Psi_T \rangle}{\langle \Phi_0 | \Psi_T \rangle} = \frac{\int f(\mathbf{R}, \infty) E_L(\mathbf{R}) d\mathbf{R}}{\int f(\mathbf{R}, \infty) d\mathbf{R}} \approx \frac{1}{M} \sum_{i=1}^M E_L(\mathbf{R}_i) = \langle E_L \rangle, \quad (3.34)$$

where E_D is the DMC energy, M is the number of configurations (walkers), and $\langle E_L \rangle$ is the average local energy. The shift energy E_T in $G_b(\mathbf{R} \leftarrow \mathbf{R}', \tau)$ is used to control the total population of walkers, such that during the simulation it is occasionally adjusted so that the overall number of walkers fluctuates around a desired mean value. Figure 3.4 illustrates an example of the behavior of the reference energy, the total population, and the average local energy. In the end, one can outline a basic version of the fixed-node DMC algorithm as follows,

1. Generate a set of walkers. (usually the initial walkers are sampled from $\Psi_T^2(\mathbf{R})$ using VMC.)
2. Choose a walker and propose a move from its old position $\mathbf{R}'$ to a new position $\mathbf{R}$ for a time step τ ,

$$\mathbf{R} = \mathbf{R}' + \chi + \tau \mathbf{v}_D(\mathbf{R}'), \quad (3.35)$$

where χ is a $3N$ -dimensional Gaussian random variable with a variance τ and zero mean.

3. Accept the move with the probability given by Equation 3.31.
4. Calculate the local energy of the walker and other quantities of interest.

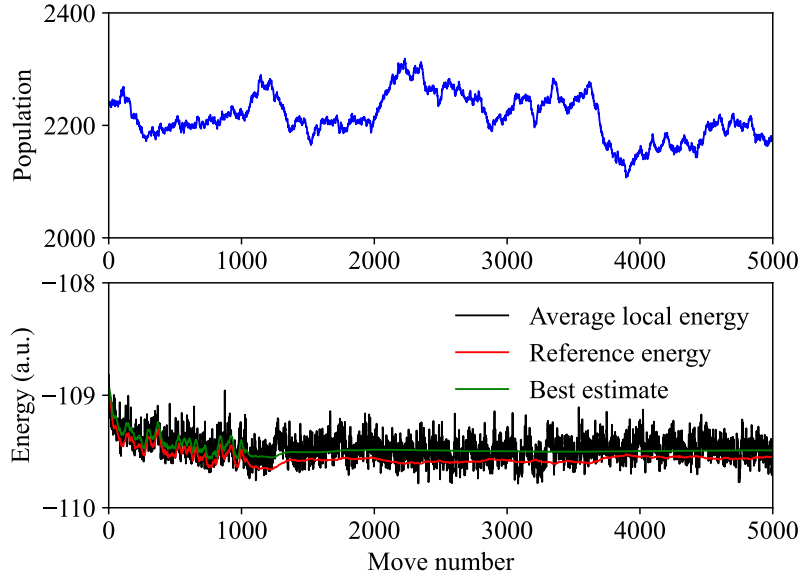


Figure 3.4: DMC data for a Nitrogen dimer N_2 at equilibrium geometry with a Jastrow-free multi-determinant wave function. The upper panel shows the fluctuations in the population of walkers for the branching process arising from simulations done via Equation 3.30. In the lower panel, the black line shows the local energy averaged over the current population of walkers, the red line shows the reference energy, and the green line is the best estimate of the DMC energy. The initial walkers come from a previous VMC calculation, and the energy decays from its initial VMC value to reach a plateau. The data before the plateau correspond to the DMC equilibration phase.

5. Calculate the number of copies or the multiplicity (M) of the walker that will continue in the evolution,

$$M = \text{int}[\eta + G_b(\mathbf{R} \leftarrow \mathbf{R}', \tau)], \quad (3.36)$$

where η is a random number between 0 and 1.

6. If $M \neq 0$, weight the local energy and other quantities of interest by M .
7. Repeat steps 2 through 7 until the times associated with different walkers have advanced the same.
8. Accumulate the local energy and other quantities of interest.

9. Update the trial energy E_T . A simple formula for adjusting E_T is⁶⁷

$$(E_T)_{new} = \frac{[(E_T)_{old} + \langle E_L \rangle]}{2}. \quad (3.37)$$

10. Renormalize the number of walkers to the initial set size by randomly eliminating or copying from the existing set of walkers.
11. Repeat steps 2 through 11 until there is no systematic trend in the cumulative local energy $\langle E_L \rangle$ and other quantities of interest. At this point the equilibration phase has been reached.
12. Repeat steps 2 through 12 until the statistical uncertainty in $\langle E_L \rangle$ and other quantities of interest has reached the desired level.

Figure 3.5 illustrates the flow chart of the fixed-node DMC algorithm.

3.4 Trial Wave Functions

Trial wave functions have a central role in both VMC and DMC, such that the statistical efficiency and accuracy of the VMC and DMC calculations are controlled by them. The accuracy of VMC results depends on the entire trial wave function, whereas the accuracy in DMC depends only on the nodal surface of the trial wave function.

3.4.1 Slater-Jastrow Wave Functions

The Jastrow correlation factor is the primary way for including electron correlation beyond the mean-field level in VMC, and this is achieved by making the wave function depend explicitly on particle separations⁷⁴. A pair of up- and down-spin determinants multiplied by a Jastrow factor is the most commonly used trial wave function in QMC,

$$\Psi_{\text{SJ}}(\mathbf{R}) = e^{J(\mathbf{R})} D(\mathbf{R}^\uparrow) D(\mathbf{R}^\downarrow), \quad (3.38)$$

where $e^{J(\mathbf{R})}$ is the Jastrow factor, and $D(\mathbf{R}^\uparrow)$ and $D(\mathbf{R}^\downarrow)$ are determinants of molecular orbitals for up- and down-spin electrons, respectively. The molecular orbitals are often obtained from density functional theory (DFT)^{6,7} or Hartree-Fock calculations. The Jastrow factor is symmetric under the interchange of electrons, and it is always positive, a

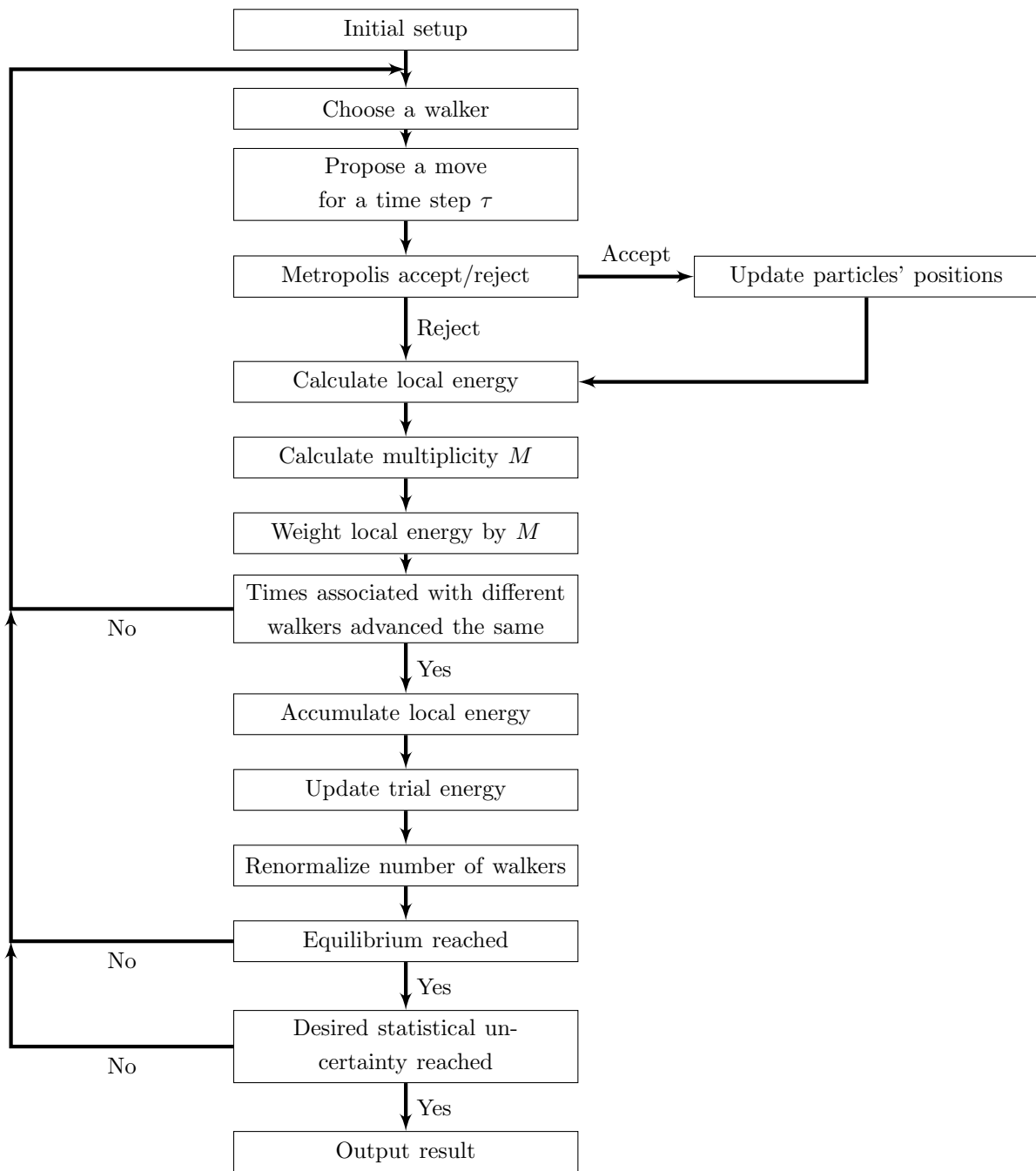


Figure 3.5: Flow chart of the fixed-node DMC algorithm.

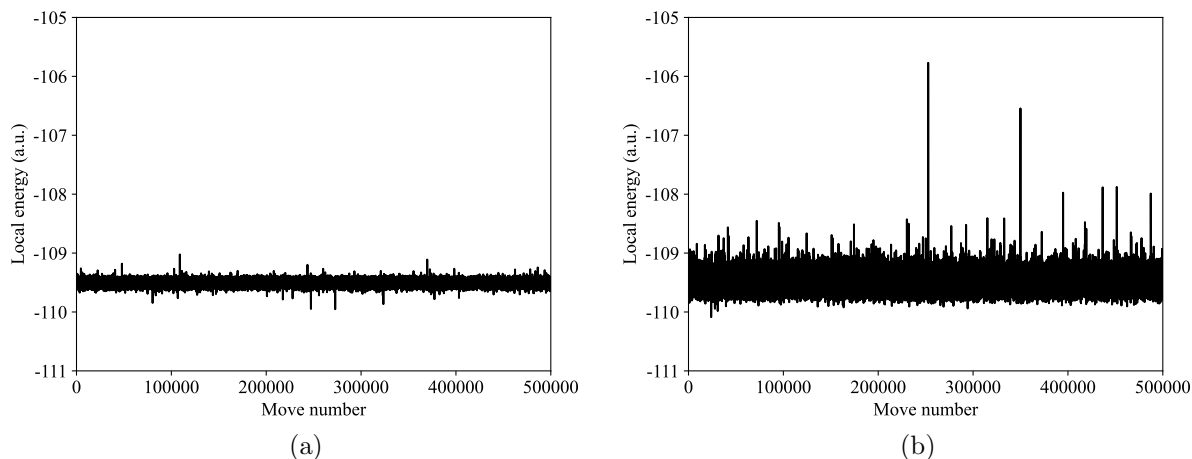


Figure 3.6: Local energy of a Nitrogen dimer N_2 at equilibrium geometry from the DMC calculation of a multi-determinant trial wave function (a) with and (b) without a Jastrow factor.

property which prevents the Jastrow factor from changing the nodal surface of the trial wave function.

The potential energy diverges whenever an electron approaches a nucleus or another electron; therefore, the exact many-electron wave function must have a cusp, such that the local kinetic energy supplies an equal and opposite divergence. This non-smoothness in the many-electron wave function can be locally resolved using the so-called Kato cusp conditions^{75,76}. Any necessary electron-nucleus cusps can be built into the molecular orbitals obtained from the DFT or HF approximation, and electron-electron cusps can be built into the Jastrow factor. The application of cusp conditions causes a reduction in both the average local energy and its variance. Figure 3.6 shows the local energies of a nitrogen dimer N_2 at equilibrium geometry obtained in two DMC calculations for a multi-determinant trial wave function with and without a Jastrow factor. The trial wave function in Figure 3.6b does not include a Jastrow factor, so the Kato cusp conditions at the electron-electron coalescence points are not satisfied. This manifests as very large positive spikes in the local energy when two electrons are close together. Figure 3.6a shows the effect of adding a Jastrow factor to the trial wave function, which removes the very large positive spikes and lowers the average local energy. After adding a Jastrow factor, some small positive and negative spikes still remain. These small spikes are due to the positive and negative diver-

gences of the local kinetic energy when electrons approach the nodal surface. Throughout this thesis, we use a Jastrow factor of the form $e^{J(\mathbf{R})}$,¹⁵

$$J(\{\mathbf{r}_i\}, \{\mathbf{r}_A\}) = \sum_{j>i}^N u(r_{ij}) + \sum_{A=1}^M \sum_{i=1}^N \chi_A(r_{iA}) + \sum_{A=1}^M \sum_{j>i}^N f_A(r_{iA}, r_{jA}, r_{ij}), \quad (3.39)$$

where u is a homogeneous and isotropic electron-electron function, χ is an isotropic electron-nucleus function centered on the nuclei, and f is an isotropic electron-electron-nucleus function, also centered on the nuclei. These functions are represented by power expansions with optimizable parameters. N and M are the number of electrons and nuclei, respectively. Electrons and nuclei are described by position vectors $\mathbf{r}_i$ and $\mathbf{r}_A$, respectively. $r_{ij} = |\mathbf{r}_{ij}| = |\mathbf{r}_i - \mathbf{r}_j|$ is the distance between i th and j th electron, $r_{iA} = |\mathbf{r}_i - \mathbf{r}_A|$ is the distance between i th electron and A th nucleus, and $r_{jA} = |\mathbf{r}_j - \mathbf{r}_A|$ is the distance between j th electron and A th nucleus. The Jastrow factor of Equation 3.39 includes two-body (e-e and e-N) and three-body (e-e-N) terms, while one can in general include higher-order terms in the Jastrow factor, such as terms involving the distances between four or more particles. The inclusion of higher order terms in the Jastrow factor reduces both the energy and its variance¹⁶.

The Jastrow factor also breaks the spin symmetry of the trial wave function⁷⁷, meaning that if the determinantal part of a Slater-Jastrow wave function is an eigenfunction of $\hat{\mathbf{S}}^2$ operator (such as configuration state functions), the inclusion of the Jastrow factor leads to a trial wave function that is not an eigenfunction of $\hat{\mathbf{S}}^2$.

3.4.2 Multi-Slater-Jastrow Wave Functions

Although the most commonly used form of a trial wave function in QMC is the Slater-Jastrow form, retrieving a greater fraction of the correlation energy requires the use of more sophisticated ansatz. This is especially the case in intrinsically multi-reference problems. A natural extension of the Slater-Jastrow wave function is the multi-Slater-Jastrow wave function,

$$\Psi_{\text{MD}}(\mathbf{R}) = e^{J(\mathbf{R})} \sum_n c_n D_n(\mathbf{R}^\uparrow) D_n(\mathbf{R}^\downarrow), \quad (3.40)$$

where c_n is the coefficient of the n th determinant pair. One advantage of this type of trial wave function is that, any (ground and excited) state of the system can be represented with a large enough number of determinants. Such wave function form however is only

useful if it can be used with limited computational cost and memory. So they are not in general suitable for large systems since the number of determinants required to retrieve a given fraction of the correlation energy increases exponentially with the system size.

3.4.3 Backflow Wave Functions

The classical backflow effect is related to the flow of a fluid around a large impurity. For example, consider a car moving with high speed. The air around the moving car is pushed out of the way and it fills in behind the car. The quantum backflow effect can then be regarded as the similar effect for quantum particles moving through a quantum fluid^{78,79}. This introduces a correlation, which can be mainly captured by the Jastrow factor. The backflow transformation also captures another contribution of this correlation, while, in contrast to the Jastrow factor, it alters the nodal surface of the trial wave function. The backflow correlations are introduced by substituting a set of coordinates $\mathbf{X}$ in place of the electron positions $\mathbf{R}$ in the Slater determinants¹⁷,

$$\Psi_{\text{BF}}(\mathbf{R}) = e^{J(\mathbf{R})} D(\mathbf{X}^\uparrow) D(\mathbf{X}^\downarrow), \quad (3.41)$$

where the new coordinate $\mathbf{X}$ for both up- and down-spin i th electron is given by

$$\mathbf{x}_i = \mathbf{r}_i + \boldsymbol{\xi}_i(\mathbf{R}), \quad (3.42)$$

where $\boldsymbol{\xi}_i(\mathbf{R})$ is the backflow displacement of the i th electron, which depends on the configuration of the whole system. For a system of N electrons and M nuclei, the backflow displacement of the i th electron would be as follows,

$$\boldsymbol{\xi}_i = \sum_{j \neq i}^N \eta_{ij} \mathbf{r}_{ij} + \sum_{A=1}^M \mu_{iA} \mathbf{r}_{iA} + \sum_{j \neq i}^N \sum_{A=1}^M (\Phi_i^{jA} \mathbf{r}_{ij} + \Theta_i^{jA} \mathbf{r}_{iA}), \quad (3.43)$$

where $\eta_{ij} = \eta(r_{ij})$ is a function of electron-electron distance, $\mu_{iA} = \mu(r_{iA})$ is a function of electron-nucleus distance, and $\Phi_i^{jA} = \Phi^A(r_{iA}, r_{jA}, r_{ij})$ and $\Theta_i^{jA} = \Theta^A(r_{iA}, r_{jA}, r_{ij})$ are electron-electron-nucleus terms to describe two-electron backflow displacements in the presence of a nearby nucleus. These functions are parametrized by power expansions with optimizable parameters. Furthermore, the use of backflow transformations usually makes the computational cost of QMC calculations a factor of N more expensive. Also, one can combine the multi-Slater-Jastrow wave function with backflow transformations to use multi-Slater-Jastrow-backflow wave function as the trial wave function.

3.5 Full Configuration Interaction Quantum Monte Carlo

Full configuration interaction is among the most robust methods to treat the electron correlations in many-electron systems; however, its application is greatly limited by the demanding computational cost and memory requirements needed for the iterative diagonalizations in the Slater determinant space. On the other hand, accurate stochastic methods such as the diffusion Monte Carlo or the Green’s function Monte Carlo suffer from the fermion sign problem, which prevents exact numerical simulation of many-electron systems. Full configuration interaction quantum Monte Carlo (FCIQMC)¹² is a stochastic projector method which unifies QMC and FCI in a way which takes advantages of both techniques. This method converges onto the FCI solution whilst requiring a fraction of the memory of an FCI calculation using the same basis set. Besides, unlike the iterative diagonalization procedure required in FCI, the FCIQMC algorithm is readily parallelizable and can be run efficiently on thousands of cores.

The FCIQMC method like all projector methods, stochastically integrates the imaginary-time Schrödinger equation (see 3.3.1). The integration, in contrast to DMC, is performed in the Slater determinant basis, which is a pure antisymmetric space. Working in the Slater determinant space does not evade the fermion sign problem, however it avoids solutions with bosonic character. The FCI wave functions do not have strictly non-negative amplitudes, instead they have a so-called sign structure, which is defined by the sign of the CI coefficients. The fermion sign problem in the Slater determinant space stems from the fact that it is impossible to have a correct sign structure for a given FCI wave function without a knowledge of its true CI coefficients. Integration in the Slater determinant space should not then be regarded as a remedy to the sign problem, but it can help to reduce the sign problem in the fermionic systems^{18,19}. Furthermore, the propagation step in the FCIQMC is based on the population dynamics (i.e., birth and death of walkers), and in common with DMC, walkers at each step define the instantaneous wave function. The key ingredient of FCIQMC is the walker annihilation step, which can be interpreted as quantum mechanical destructive interference. Each walker during the simulation carries a positive or negative sign, such that due to the annihilation step, each pair of walkers having opposite signs and occupying the same determinant are removed from the simulation.

To begin with, the imaginary-time Schrödinger equation is rewritten as

$$\hat{H}\Psi(\tau) = -\frac{\partial\Psi(\tau)}{\partial\tau}, \quad (3.44)$$

where $\Psi(\tau)$ at any imaginary time τ can be given by

$$\Psi(\tau) = \sum_i c_i(\tau) |D_i\rangle, \quad (3.45)$$

where $c_i(\tau)$ are time-dependent CI coefficients. Let us now define a new matrix K ,

$$K_{ij} = \langle D_i | \hat{K} | D_j \rangle = \langle D_i | \hat{H} | D_j \rangle - E_{\text{HF}} \delta_{ij}, \quad (3.46)$$

whose matrix elements are those of the Hamiltonian $\hat{H}$, except for the diagonal elements which are subtracted from the Hartree-Fock energy E_{HF} . Consequently, the lowest energy eigenvalue of $\hat{K}$ would be $E_0^{\text{FCI}} - E_{\text{HF}}$, which is the correlation energy E_{corr} of the system. Next, the following equations can be obtained by substituting Equations 3.45 and 3.46 into 3.44,

$$\begin{aligned} -\frac{dc_i}{d\tau} &= \sum_j (K_{ij} - S \delta_{ij}) c_j \\ &= (K_{ii} - S) c_i + \sum_{j \neq i} K_{ij} c_j, \end{aligned} \quad (3.47)$$

where S is an arbitrary energy shift subtracted from the diagonal terms. If a vector $\mathbf{C}$, which is made up of the CI coefficients c_i , satisfies the following equations

$$\sum_j K_{ij} c_j = S c_i, \quad (3.48)$$

then according to Equations 3.47, $d\mathbf{C}/d\tau = 0$, making the vector $\mathbf{C}$ stationary. Based on the above equations, the vector $\mathbf{C}$ is also an eigenstate of $\hat{K}$ with eigenvalue S . Given that, if the energy shift S equals the correlation energy E_{corr} , the stationary state is then the ground state of the Hamiltonian $\hat{H}$. The instantaneous wave function is also represented by the CI coefficients $\{c_i\}$, and the i th CI coefficient is itself represented by a set of walkers occupying the i th Slater determinant,

$$c_i \propto \sum_{\alpha} s_{\alpha} \delta_{i,i_{\alpha}} = N_i, \quad (3.49)$$

where the sum is over all walkers, each carrying a sign $s_{\alpha} = \pm 1$ and occupying determinant i_{α} . N_i is then the signed sum of walkers on determinant i , which can be positive or negative. As mentioned before, FCIQMC makes use of the population dynamics to propagate the walkers throughout the discrete Slater determinant space. It follows that the entire space of (symmetry-related) determinants can eventually be reached starting from any one determinant. The population dynamics algorithm is used to simulate the set of coupled differential equations of 3.47 at each time step $\delta\tau$ as follows,

1. Each walker α , located on the determinant D_{i_α} , selects a connected determinant D_j with normalized probability $p_{gen}(j|i_\alpha)$, and spawns a progeny there with probability $p_s(j|i_\alpha) = \delta\tau|K_{i_\alpha j}|/p_{gen}(j|i_\alpha)$. The connected determinant D_j is restricted to be one of the single or double excitations of D_{i_α} (see Appendix A). The sign of the spawned walker is also the same as the parent if $K_{i_\alpha j} < 0$, and opposite to the parent otherwise.
2. Each parent walker attempts to die or clone itself with probability $p_d(i_\alpha) = \delta\tau(K_{i_\alpha i_\alpha} - S)$, such that it dies if $p_d > 0$ with probability p_d , and it is cloned if $p_d < 0$ with probability $|p_d|$.
3. Each pair of walkers of opposite sign occupying the same determinant is removed from the list of walkers.

A simulation based on the above algorithm converges to the FCI solution (i.e. FCI wave function and energy). Besides, one can reduce the total number of walkers N_w required to achieve convergence, and hence reduce the computational cost and memory requirement, by using the initiator approximation¹³. This approximation defines a survival criterion for the newly spawned walkers. The progeny of walkers spawned from a predefined set of determinants (known as initiator determinants) are able to survive, while those from non-initiator determinants can survive only if they are spawned onto determinants which are already occupied. An exception to this rule is when two non-initiator determinants spawn walkers of the same sign onto a third unoccupied determinant in one iteration. This is considered as a sign-coherent event and the two newly spawned walkers can survive. The initial set of initiators is defined according to a complete-active-space criterion (CAS) which can be constructed from a subset of orbitals. However, the set of initiators during the simulation is updated dynamically, such that any non-initiator determinant acquiring a population of walkers which exceeds a user-defined value N_{init} becomes an initiator, and stays so as long as its occupation number remains above N_{init} . The FCIQMC simulation consists of two phases,

1. Fixed-shift mode: during this phase, the shift S is fixed at a constant value above the correlation energy E_{corr} . This results in an exponential growth of walkers, whose rate depends on the shift value, the time step, and the correlation energy.
2. Fixed- N_w mode: when a desired number of walkers is reached, the shift S varies to attempt to keep this number constant. Finally, the shift will fluctuate about the correct correlation energy, when the walkers are distributed according to the FCI ground-state wave function (for large enough N_w).

To start an FCIQMC simulation, a small number of walkers is placed on the HF determinant. Then, the population dynamics is used to grow and propagate the walkers during the two phases (fixed-shift and fixed- N_w). The value of the shift eventually will give the correlation energy of the system. Moreover, as FCIQMC is a projector method, one can use the projected energy estimator to have another measure of the correlation energy.

$$\begin{aligned}
E(\tau) &= \frac{\langle D_0 | \hat{H} | \Psi(\tau) \rangle}{\langle D_0 | \Psi(\tau) \rangle} \\
&= E_{\text{HF}} + \sum_{i \neq 0} \langle D_i | \hat{H} | D_0 \rangle \frac{c_i(\tau)}{c_0(\tau)} = E_{\text{HF}} + \sum_{i \neq 0} \langle D_i | \hat{H} | D_0 \rangle \frac{N_i(\tau)}{N_0(\tau)},
\end{aligned} \tag{3.50}$$

where D_0 is the HF determinant and $N_i(\tau)$ is the time-dependent number of walkers on determinant D_i defined by Equation 3.49. Note that this summation runs over all single and double excitations, as those are the only determinants connected to D_0 . In the long-time limit, the projected energy $E(\tau)$ eventually converges onto the FCI ground-state energy as the ratio N_i/N_0 tends to c_i/c_0 of the FCI wave function. Figure 3.7 illustrates an FCIQMC simulation for a carbon dimer C_2 at equilibrium geometry in a cc-pCVDZ basis set. One can readily differentiate the two different phases in the plots. Also as shown, in the long-time limit, the projected energy matches the energy shift, introducing a measure for the correlation energy. In this thesis, all FCIQMC calculations are done using the NECI code^{12,80}, which is developed and maintained by the Alavi group.

As an example, we perform FCIQMC calculations for the carbon dimer in Dunning basis sets, cc-pCVnZ ($n=\text{D, T, Q}$), and then extrapolate the correlation energy to the complete-basis-set (CBS) limit using the following fitting function²¹.

$$E_{\text{corr}}(M) = E_{\text{corr}}(\infty) + aM^{-1} + bM^{-2}, \tag{3.51}$$

where M is the basis-set size, and $E_{\text{corr}}(\infty)$, a , and b are fit parameters. The CBS extrapolation, as shown in Figure 3.8, returns a total energy of -75.9278(12) Ha. The obtained energy slightly differs from the exact ground-state energy of -75.9265 Ha⁵. This error mainly comes from the initiator approximation, namely the initiator error, which can be reduced by using larger number of walkers. We will introduce a new extrapolation scheme to get the infinite-walker limits in chapter 6.

3.5.1 The Transcorrelated FCIQMC

Electronic correlation can be categorized broadly into static and dynamic correlation. In conventional descriptions of the many-body wave function, both types of correlation are

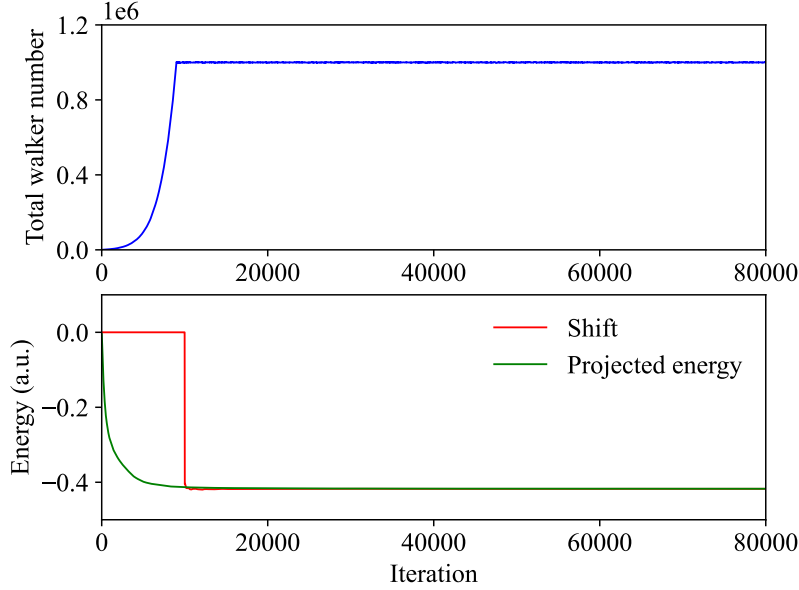


Figure 3.7: The upper plot shows the walker growth for a carbon dimer C_2 at equilibrium geometry in a cc-pCVDZ basis set in the fixed-shift mode ($S = 0$). After reaching a target population of 10^7 walkers, the shift, in the lower plot, starts to vary to keep the number of walkers constant. The projected energy is also shown in the lower plot. The shift value which stabilizes the walker growth is in agreement with the projected energy.

typically treated using a linear expansion of Slater determinants. While this configuration approach proves effective for handling static correlation in a natural and efficient manner, it falls short in efficiently addressing dynamic correlation, often leading to slow convergence towards the complete basis limit. The primary challenge arises from the singular nature of the Coulomb potential, which at particle coalescence points introduces non-smoothness (cusps) into the many-body wave function. These irregularities cannot be adequately approximated using orbital product expansions, which are inefficient at describing dynamic correlation. As stated in 3.4.1, the non-smooth behavior of the many-body wave function can be locally resolved using the Kato cusp conditions. One way of incorporating the cusp conditions is to use the Jastrow ansatz as follows,

$$\Psi = e^J \Phi, \quad (3.52)$$

where J is a symmetric correlation factor and Φ is an antisymmetric function, which can be a single Slater determinant or a linear combination of them. Given that J effectively

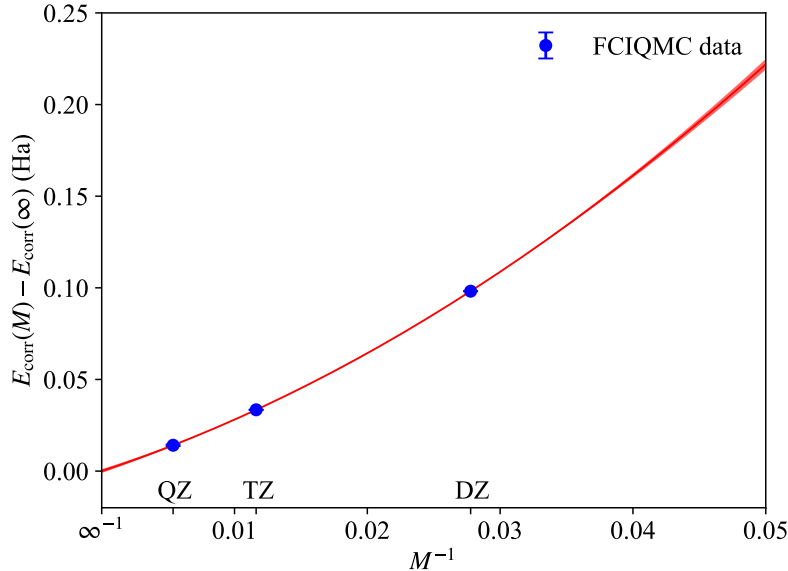


Figure 3.8: Finite-basis-set error in the FCIQMC calculations for the correlation energy of a carbon dimer C_2 at equilibrium geometry as a function of the inverse basis-set size M^{-1} . The data correspond to calculations in Dunning basis sets, cc-pCV n Z ($n=\text{D, T, Q}$). The line width indicates the statistical uncertainty in the fit.

handles the majority of the non-analytic aspects, Φ is likely to exhibit a rather smooth behavior, allowing for accurate representations with smaller basis sets. To handle such an ansatz efficiently, Boys and Handy^{81–89} proposed a relatively straightforward approach known as the transcorrelated (TC) method. This method involves utilizing a similarity transformation, $e^{-J}\hat{H}e^J$, to eliminate the exponential Jastrow factor from the equations involved. Due to the non-Hermitian nature of the resulting effective Hamiltonian, the energy is not prevented from going below the exact value. The absence of a variational bound has been regarded as a big concern, significantly limiting the widespread adoption of the TC method for an extended period of time. The method was later successfully combined with quantum chemistry methods, such as perturbation theory and the linearized coupled cluster method^{90–92}. Incorporating the TC method into QMC approaches, such as VMC^{93,94} and FCIQMC^{95–98} revived interest in the approach for the study of strongly correlated systems.

Substituting the Jastrow ansatz of Equation 3.52 into the Schrödinger equation, $\hat{H}\Psi =$

$E\Psi$, gives the similarity-transformed Schrödinger equation as

$$\hat{H}_{\text{TC}}\Phi = E\Phi, \quad (3.53)$$

where the transcorrelated Hamiltonian is

$$\hat{H}_{\text{TC}} = e^{-J}\hat{H}e^J, \quad (3.54)$$

and Φ is the right-eigenvector of $\hat{H}_{\text{TC}}$. By employing the Baker-Campbell-Hausdorff expansion, the TC Hamiltonian can be precisely evaluated, and for correlation factors that only depend on electronic positions, the truncation of this expansion occurs exactly at the second order,

$$\hat{H}_{\text{TC}} = e^{-J}\hat{H}e^J = \hat{H} + [\hat{H}, J] + \frac{1}{2}[[\hat{H}, J], J]. \quad (3.55)$$

As previously stated, the FCIQMC method relies on Monte Carlo simulations that dynamically evolve the many-body wave function with imaginary time,

$$\Psi(\tau) = e^{-\tau(\hat{H}-E_0)}\Psi(\tau=0), \quad (3.56)$$

which, as all projector methods, leads to the ground-state wave function in the long-time limit $\Psi_0 = \Psi(\tau \rightarrow \infty)$. On the other hand, the time evolution of the wave function for the Jastrow ansatz is expressed as

$$\Psi(\tau) = e^J\Phi(\tau), \quad (3.57)$$

which can be substituted into Equation 3.56 to yield

$$\begin{aligned} \Phi(\tau) &= e^{-J}e^{-\tau(\hat{H}-E_0)}e^J\Phi(\tau=0) \\ &= \lim_{M \rightarrow \infty} e^{-J} \left[1 - \frac{\tau(\hat{H} - E_0)}{M} \right]^M e^J \Phi(\tau=0) \\ &= \lim_{M \rightarrow \infty} \left[e^{-J} \left[1 - \frac{\tau(\hat{H} - E_0)}{M} \right] e^J \right]^M \Phi(\tau=0) \\ &= \lim_{M \rightarrow \infty} \left[1 - \frac{\tau(\hat{H}_{\text{TC}} - E_0)}{M} \right]^M \Phi(\tau=0) \\ &= e^{-\tau(\hat{H}_{\text{TC}} - E_0)}\Phi(\tau=0). \end{aligned} \quad (3.58)$$

The correspondence between Equations 3.56 and 3.58 demonstrates that the FCIQMC method can handle properly the non-Hermiticity arising from these types of similarity transformations⁹⁵. Therefore, the FCIQMC method can be directly applied to the TC Hamiltonian $\hat{H}_{\text{TC}}$, offering significant enhancements in basis-set convergence⁹⁶ and in compactness of CI expansions, for example in the case of strongly correlated two-dimensional Hubbard models⁹⁷. The only technical modification required in comparison with a non-TC FCIQMC implementation lies in handling non-Hermitian operators, which results in non-symmetric matrix elements.

For a symmetric correlation factor over electron pairs,

$$J = \sum_{i < j} u(\mathbf{r}_i, \mathbf{r}_j), \quad (3.59)$$

where $u(\mathbf{r}_i, \mathbf{r}_j) = u(\mathbf{r}_j, \mathbf{r}_i)$, the explicit form of $\hat{H}_{\text{TC}}$ including additional two- and three-body terms is

$$\begin{aligned} \hat{H}_{\text{TC}} &= \hat{H} - \sum_i \left(\frac{1}{2} \nabla_i^2 J + (\nabla_i J) \cdot \nabla_i + \frac{1}{2} (\nabla_i J)^2 \right) \\ &= \hat{H} - \sum_{i < j} \hat{K}(\mathbf{r}_i, \mathbf{r}_j) - \sum_{i < j < k} \hat{L}(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k), \end{aligned} \quad (3.60)$$

where

$$\begin{aligned} \hat{K}(\mathbf{r}_i, \mathbf{r}_j) &= \frac{1}{2} [\nabla_i^2 u(\mathbf{r}_i, \mathbf{r}_j) + \nabla_j^2 u(\mathbf{r}_i, \mathbf{r}_j) + |\nabla_i u(\mathbf{r}_i, \mathbf{r}_j)|^2 + |\nabla_j u(\mathbf{r}_j, \mathbf{r}_i)|^2] \\ &\quad + \nabla_i u(\mathbf{r}_i, \mathbf{r}_j) \cdot \nabla_i + \nabla_j u(\mathbf{r}_i, \mathbf{r}_j) \cdot \nabla_j, \\ \hat{L}(\mathbf{r}_i, \mathbf{r}_j, \mathbf{r}_k) &= \nabla_i u(\mathbf{r}_i, \mathbf{r}_j) \cdot \nabla_i u(\mathbf{r}_i, \mathbf{r}_k) + \nabla_j u(\mathbf{r}_j, \mathbf{r}_i) \cdot \nabla_j u(\mathbf{r}_j, \mathbf{r}_k) \\ &\quad + \nabla_k u(\mathbf{r}_k, \mathbf{r}_i) \cdot \nabla_k u(\mathbf{r}_k, \mathbf{r}_j). \end{aligned} \quad (3.61)$$

With this first-quantized Hamiltonian at hand, it becomes possible to formulate a second-quantized Hamiltonian for a given set of orthonormal spatial orbitals $\{\psi_1, \dots, \psi_{n_{\text{orb}}}\}$, with corresponding fermionic spin- $\frac{1}{2}$ creation (annihilation) operators $\hat{a}_{p\sigma}^\dagger$ ($\hat{a}_{p\sigma}$),

$$\begin{aligned} \hat{H}_{\text{TC}} &= \sum_{pq\sigma} h_q^p \hat{a}_{p\sigma}^\dagger \hat{a}_{q\sigma} + \frac{1}{2} \sum_{pqrs} (V_{rs}^{pq} - K_{rs}^{pq}) \sum_{\sigma\tau} \hat{a}_{p\sigma}^\dagger \hat{a}_{q\tau}^\dagger \hat{a}_{s\tau} \hat{a}_{r\sigma} \\ &\quad - \frac{1}{6} \sum_{pqrstu} L_{stu}^{pqr} \sum_{\sigma\tau\lambda} \hat{a}_{p\sigma}^\dagger \hat{a}_{q\tau}^\dagger \hat{a}_{r\lambda}^\dagger \hat{a}_{u\lambda} \hat{a}_{t\tau} \hat{a}_{s\sigma}, \end{aligned} \quad (3.62)$$

where $h_q^p = -\frac{1}{2}\langle\psi_p|\hat{h}|\psi_q\rangle$ and $V_{rs}^{pq} = \langle\psi_p\psi_q|r_{12}^{-1}|\psi_r\psi_s\rangle$ are the one- and two-body integrals of the original Hamiltonian, and $K_{rs}^{pq} = \langle\psi_p\psi_q|\hat{K}|\psi_r\psi_s\rangle$ and $L_{stu}^{pqr} = \langle\psi_p\psi_q\psi_r|\hat{L}|\psi_s\psi_t\psi_u\rangle$ are the corresponding terms arising from the similarity transformation⁹⁶. As an example, we calculate the total energy of an unpolarized 14-electron gas at $r_s = 0.5$ with 1020 spin orbital basis functions. The calculations are done for different number of walkers n_w , using the FCIQMC method and the TC-FCIQMC method containing up to two-body operators (refer to Equations 3.60 and 3.61). As shown in Figure 3.9, the TC-FCIQMC technique efficiently captures a larger portion of the correlation energy. Consequently, the transcorrelated method elevates the effective level of theory of the underlying correlation approach, serving as a significant incentive to develop this methodology with the ultimate goal of investigating strongly correlated *ab initio* systems. To this end, a novel approach has been recently developed to optimize highly flexible Jastrow factors customized for individual systems for use in the transcorrelated method⁹⁸. This technique offers extensions both within the domain of TC-FCIQMC and in other approaches, such as TC coupled-cluster theory^{99–102}, TC density matrix renormalization group^{103,104}, TC selected-CI approaches^{105,106}, and also for reducing the resources required in a quantum computing setting^{107,108}.

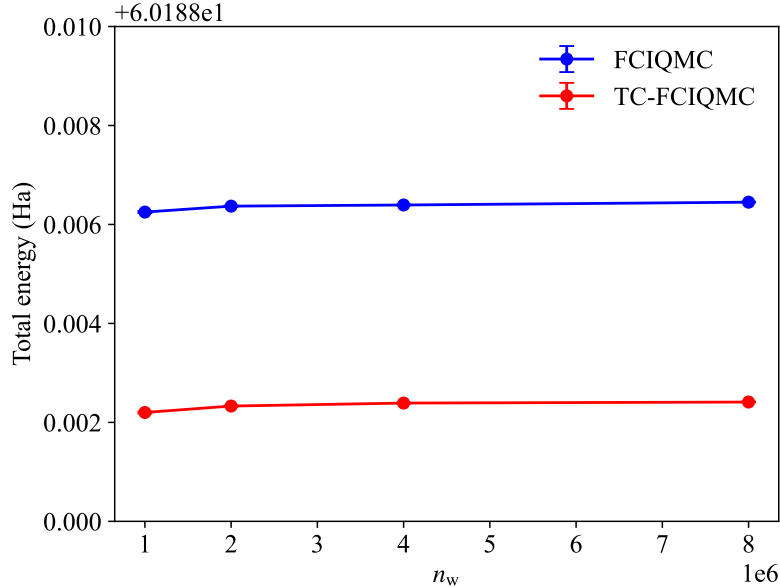


Figure 3.9: Total energy of an unpolarized 14-electron gas at $r_s = 0.5$ obtained with FCIQMC and TC-FCIQMC methods as a function of the number of walkers.

Part III

Results and Applications

Chapter 4

Accurate Energies from Intersecting Extrapolations of Continuum Quantum Monte Carlo Data

We present an extrapolation method that yields very accurate total and relative energies from variational and diffusion quantum Monte Carlo (VMC and DMC) results. For a trial wave function consisting of a small configuration interaction (CI) wave function obtained from full CI quantum Monte Carlo and reoptimized in the presence of a Jastrow factor and an optional backflow transformation, we find that the VMC and DMC energies are smooth functions of the sum of the squared coefficients of the initial CI wave function, and that quadratic extrapolations of the VMC and DMC energies intersect within uncertainty of the exact total energy. With adequate statistical treatment of quasi-random fluctuations, the *extrapolated intersection with polynomials of order two* (XSPOT) method is shown to yield results in agreement with benchmark-quality total and relative energies for the C₂, N₂, H₂O, and CO₂ molecules, as well as for the C₂ molecule in its first electronic singlet excited state, using only small CI expansion sizes.ⁱ

4.1 Introduction

Quantum Monte Carlo (QMC) methods are a broad family of stochastic wave-function-based techniques that accurately approximate the solution of the Schrödinger equation of

ⁱThis chapter is based on Ref. 109, written in collaboration with Ali Alavi and Pablo López Ríos.

an electronic system. The variational quantum Monte Carlo (VMC) method^{10,14} obtains expectation values corresponding to an analytic trial wave function $\Psi_T(\mathbf{R})$ in real space and provides a framework for optimizing wave function parameters,^{110,111} such as those in the multideterminant-Jastrow-backflow form,

$$\Psi_T(\mathbf{R}) = e^{J(\mathbf{R})} \sum_I^M c_I D_I[\mathbf{X}(\mathbf{R})], \quad (4.1)$$

where $D_I(\mathbf{R})$ are M Slater determinants, $e^{J(\mathbf{R})}$ is a Jastrow correlation factor,^{15,16} and $\mathbf{X}(\mathbf{R})$ are backflow-transformed electronic coordinates.¹⁷ Diffusion quantum Monte Carlo (DMC)^{11,14} is a real-space projection method which recovers the lowest-energy solution $\Phi(\mathbf{R})$ of the Schrödinger equation compatible with the fixed-node condition that $\Phi(\mathbf{R})\Psi_T(\mathbf{R})$ be nonnegative everywhere. We refer to VMC and DMC collectively as continuum QMC (cQMC) methods.

In the configuration interaction (CI) ansatz the solution of the Schrödinger equation in a given orbital basis is expressed as

$$|\Psi\rangle = \sum_I c_I |D_I\rangle, \quad (4.2)$$

where $\{|D_I\rangle\}$ are all possible Slater determinants that can be constructed in the orbital basis. Equation 4.2 is exact in the limit of an infinite expansion using a complete (infinite) orbital basis, but in practical methods finite expansions and bases are used. Full CI quantum Monte Carlo (FCIQMC)^{12,13} is a second-quantized projection technique in which random walkers sample the discrete Hilbert space defined by $\{|D_I\rangle\}$ in order to determine approximate values of the CI coefficients $\{c_I\}$.

The computational cost of cQMC calculations scales polynomially with system size N , typically as N^2 – N^4 , and the quality of the resulting energies depends on the accuracy of the trial wave function for VMC, and on the accurate location of its nodes for DMC. The cQMC methods excel at describing explicit dynamic and long-ranged correlations, but the error incurred by the fixed-node approximation is often significant. By contrast, FCIQMC is formally an exponentially-scaling method which trivially captures static correlations, but requires a very large number of walkers to provide a good description of dynamic correlations. The complementary nature of the strengths of cQMC and FCIQMC makes combining these methods highly desirable. Several ways of combining cQMC and FCIQMC have been presented in the literature, such as using DMC to assist in the extrapolation to the thermodynamic limit of FCIQMC energies of the electron gas,²¹ or using VMC-optimized Jastrow factors in FCIQMC with the transcorrelated method.^{96,98} Here we focus

on the use of selected CI wave functions generated with FCIQMC to construct multi-determinantal trial wave functions for cQMC calculations.

Multideterminant expansions have been used in cQMC calculations for decades. In the case of atomic and molecular systems, various successful applications have been reported so far, including ground-state energy calculations,^{112–117} excitation energies,^{118–121} and geometry optimizations.^{120,122} The use of truncated CI expansions in cQMC presents the problem that no reliable criteria exist to truncate wave functions for different systems in a consistent manner, resulting in energy differences of questionable accuracy. One possible approach is to use extremely large multi-determinantal wave functions,^{117–125} under the expectation that the fixed-node error in the total energies will become smaller than the target error. While algorithmic developments have vastly reduced the computational cost associated with the use of multi-determinantal wave functions in cQMC,^{123,126–129} this remains an expensive choice. Using trial wave functions without a Jastrow factor reduces the nominal computational burden^{118,119,125} at the cost of losing the accurate, compact description of dynamic correlation afforded by fully-optimized trial wave functions. By including explicit correlations, we are able to explore the use of relatively small multi-determinantal wave functions to perform an extrapolation of the cQMC total energy to the full-CI, complete orbital-basis limit. We test our method on a variety of molecular systems, obtaining total and relative energies within uncertainty of benchmark-quality results from the literature.

The rest of this chapter is structured as follows. In Section 4.2, we present the methodological details of our extrapolation method, which we illustrate with calculations of the carbon dimer and the water molecule. We then apply our method to several atomic and molecular systems, and we report the results in Section 4.3.

4.2 Methodology

Let M_{gen} be the number of determinants occupied at a given point in an equilibrated FCIQMC calculation, representing the CI wave function

$$|\Psi_{M_{\text{gen}}}\rangle = \sum_{I=1}^{M_{\text{gen}}} c_I |D_I\rangle, \quad (4.3)$$

where c_I is obtained as the sum of the signed weights of the walkers occupying the I th determinant. The values of the first few coefficients $\{c_I\}_{I \ll M_{\text{gen}}}$ converge relatively quickly in FCIQMC calculations and can be expected to be reasonably close to their full CI (FCI) values. This makes FCIQMC an ideal method for quickly generating good-quality selected

CI wave functions of moderate sizes – studying the suitability of other CI solvers for this purpose is beyond the scope of this work. Let us consider the wave function obtained by truncating Equation 4.3 to size $M \ll M_{\text{gen}}$. The sum of the squares of the coefficients of the resulting wave function relative to that at size M_{gen} is

$$w^2 = \frac{\sum_{I=1}^M c_I^2}{\sum_{I=1}^{M_{\text{gen}}} c_I^2}, \quad (4.4)$$

which goes to 1 as $M \rightarrow M_{\text{gen}}$. This CI wave function of size M can be combined with a Jastrow factor, and optionally with a backflow transformation, to produce a multideterminant-Jastrow(-backflow) trial wave function for cQMC, as per Equation 4.1. The wave function parameters can be optimized in the context of VMC, producing an optimized trial wave function with which we compute VMC and DMC energies. Repeating this procedure by truncating the original CI wave function to different sizes produces a set of VMC and DMC energies that can be plotted as a function of w .

Plots of this kind, albeit using other CI solvers, can be found in the literature; see Figure 3 of Ref. 111 or Figure 4 of Ref. 127, for example. The present work is in fact inspired by the observation that the VMC and DMC curves in these plots appear to be smooth and would seem to be about to intersect just off the right-hand side of the graph. In Figure 4.1, we plot the VMC and DMC energies we obtain for the ground state of the all-electron C_2 molecule, see Table 4.1, using Hartree-Fock orbitals expanded in the cc-pCVTZ basis set,^{40,41} along with quadratic fits to the data of the form

$$E(w) = a + bw + cw^2, \quad (4.5)$$

where a , b , and c are fit parameters. In this simple example, the fits to the VMC and DMC data intersect at $w = 1.031$, corresponding to a total energy of -75.9287 Ha, not far off the exact non-relativistic total energy estimate of -75.9265 Ha given in Ref. 5. We refer to this way of estimating the total energy of a system as the *extrapolated intersection with polynomials of order two* (XSPOT) method.

In what follows we develop the methodology to enable the application of the XSPOT method in practice, and we demonstrate the different aspects of the approach using as test systems the C, N, and O atoms, the ground-state C_2 , N_2 , H_2O , and CO_2 molecules, and the C_2 molecule in its lowest-lying singlet electronic excited state $B^1\Delta_g$. All of these systems are simulated as all-electron, both in the sense that no effective-core potentials are used and that excitations from “core” orbitals are allowed in the CI wave function. In Table 4.1, we detail the electronic states and geometries we have used for these systems.

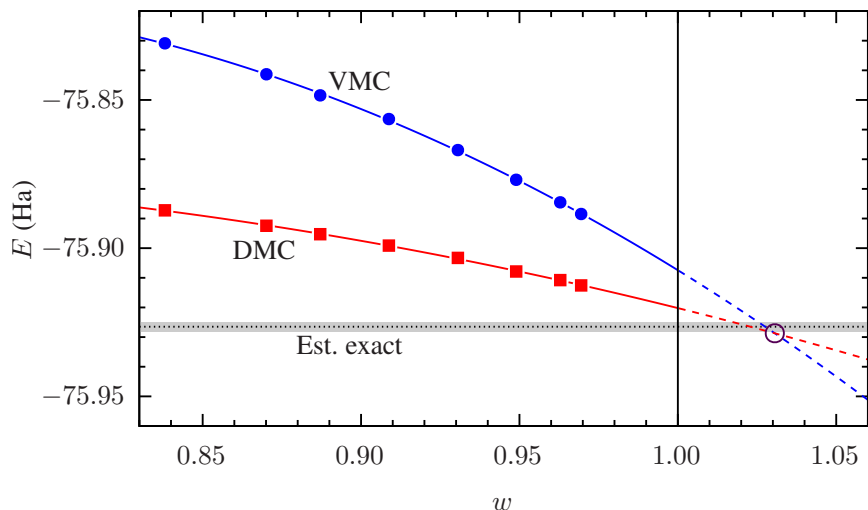


Figure 4.1: VMC and DMC total ground-state energy of the carbon dimer using multi-determinant-Jastrow trial wave functions as a function of w , using Hartree-Fock orbitals expanded in the cc-pCVTZ basis set. Quadratic fits to the data are extended beyond $w = 1$ to show their intersection, which is in good agreement with the estimated exact non-relativistic total energy of the system.⁵

4.2.1 Theoretical Justification

The extrapolation shown in Figure 4.1 might seem simplistic from a quantum chemical perspective, given that all calculations involved have been performed with the same, finite orbital basis, so one would expect an orbital-basis dependent result which should itself be extrapolated to the complete-basis limit. For instance, the FCIQMC energy tends to the basis-set dependent FCI limit as the number of walkers tends to infinity, and this must in turn be extrapolated to the basis-set limit in order to obtain the exact energy of the system.

The XSPOT extrapolation method can be easily justified in the hypothetical case of using an infinite, “complete” orbital basis. The FCIQMC energy with this “complete” orbital basis would tend to the exact total energy of the system E_0 in the infinite walker-number limit, and the sum of the squared CI coefficients would also tend to that of the exact wave function, w_0 . The exact wave function has no dynamic correlation left to recover, so the Jastrow factor and backflow displacement in the cQMC trial wave function would optimize to zero, and the VMC and DMC energies would both coincide with E_0 . At finite

System	Geometry ^{1,5}	State
C		3P
N		4S
O		3P
C ₂	$r_{CC} = 1.2425 \text{ \AA}$	$^1\Sigma_g^+$
C ₂ [*]	$r_{CC} = 1.2425 \text{ \AA}$	$B^1\Delta_g$
H ₂ O	$r_{OH} = 0.9572 \text{ \AA}$ $\angle_{HOH} = 104.5^\circ$	1A_1
N ₂	$r_{NN} = 1.0977 \text{ \AA}$	$^1\Sigma_g^+$
CO ₂	$r_{CO} = 1.1600 \text{ \AA}$ $\angle_{OCO} = 180^\circ$	$^1\Sigma_g^+$

Table 4.1: Atoms and molecules considered in this work, along with their electronic states and geometries.

expansion sizes, $w < w_0$, the VMC and DMC methods yield variational energies satisfying $E_{\text{VMC}} \geq E_{\text{DMC}} \geq E_0$, which, assuming these to be smooth functions of w , validates the XSPOT method with a “complete” orbital basis.

We note that in a truncated CI wave function, the infinite, “complete” orbital basis is effectively a finite orbital basis, since a finite number of determinants can only contain a finite number of distinct orbitals. Conversely, a sufficiently small selected CI wave function with a finite orbital basis is indistinguishable from a CI wave function of the same size with the “complete” orbital basis – assuming the finite basis contains the first few orbitals in the “complete” orbital basis. In the context of this simplified picture, CI wave functions that use all the orbitals in a finite basis can be regarded to be “exhausting” the basis, which occurs at values of w above an onset $w_{\text{exh.}}$. The cQMC energies obtained for $w > w_{\text{exh.}}$ can be expected to hit a plateau and tend to an orbital-basis dependent limit with increasing w . In practice, orbital bases such as natural orbitals can be constructed so as to compactly describe the system with fewer orbitals, which has the side effect of triggering the plateauing effects of exhaustion before having used all the orbitals in the basis. We discuss this aspect further in Section 4.2.3.

Finite-orbital-basis FCIQMC and cQMC calculations performed at $w < w_{\text{exh.}}$ behave as if one were using the “complete” orbital basis. Therefore it is legitimate to expect that the extrapolation of quadratic fits to these VMC and DMC data intersect at $w = w_0$ and $E = E_0$, provided that the VMC and DMC energies are smooth functions of w representable

by a second-order polynomial below $w_{\text{exh.}}$. This can be expected to be true in practice for multi-determinantal wave functions with w above a certain threshold $w_{\text{h.o.}}$ below which higher-order polynomials would be needed.

Note that the initial FCIQMC wave function with M_{gen} determinants is not required to be below the exhaustion limit since it simply serves to construct selected CI wave functions of size $M \ll M_{\text{gen}}$, which are required to be below the exhaustion limit, and to define the arbitrary point at which $w = 1$ in the plots; $w = 1$ has no special significance in this method. Table 4.2 shows the degree of orbital-basis exhaustion exhibited by CI wave functions of the same size using Hartree-Fock orbitals expanded in four different basis sets in the cc-pV x Z and cc-pCV x Z families^{40,41} for the all-electron carbon dimer.

Basis set	Exhaustion
cc-pVDZ	28/28 = 100%
cc-pCVDZ	32/36 = 89%
cc-pVTZ	38/60 = 82%
cc-pCVTZ	40/86 = 47%

Table 4.2: Fraction of spatial orbitals in the Hartree-Fock orbital basis expanded in each of four basis sets that appear in the first 300 configuration state functions of the full CI wave function for the ground state of the carbon dimer.

Based on these numbers, we use the cc-pCVTZ basis in this work to ensure we have enough leeway to increase the multi-determinant wave function size before hitting the exhaustion limit; this is further discussed in Section 4.3. In our calculations we choose CI wave function sizes so that the points are more or less evenly spaced in the w axis, and we make sure that different points correspond to wave functions containing a different number of distinct spatial orbitals so as to capture the effect of simultaneously growing the CI expansion and the orbital basis.

4.2.2 Obtaining Statistically Meaningful Results

In practice, cQMC energies do not exactly follow smooth curves, but it is reasonable to assume that a smooth underlying trend $E(w)$ exists, and that the cQMC energy E_i deviates from it by a quasi-random amount q_i . Considering also the statistical uncertainty ΔE_i , the i th point in a set of cQMC energies can then be modelled as

$$E_i = E(w_i) + q_i + \zeta_i \Delta E_i, \quad (4.6)$$

where ζ_i is a random number drawn from the standard normal distribution. In order to make this generic model for the quasi-random fluctuations useful in practice, we make the approximation that $q_i = \xi_i \alpha$, where ξ_i is a random number drawn from the standard normal distribution and α is a constant amplitude independent of w . Figure 4.2 illustrates our model of cQMC energy data as a function of w .

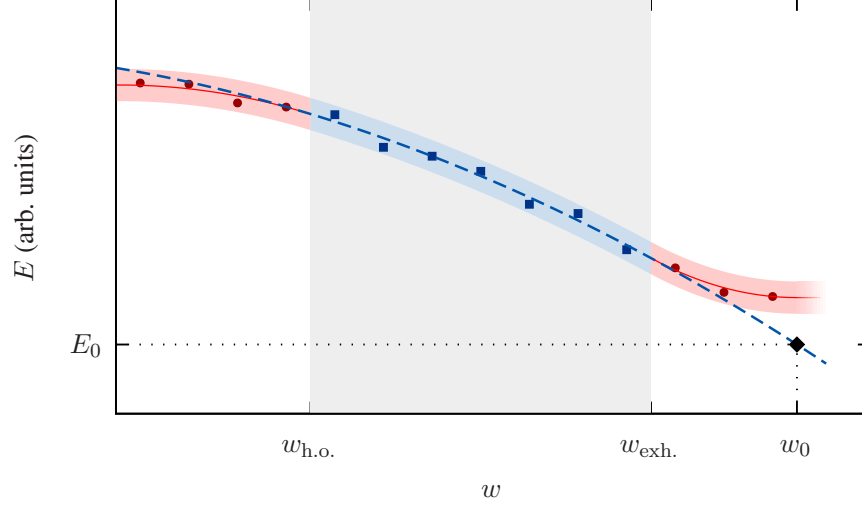


Figure 4.2: Illustration of the expected behavior of cQMC energies as a function of w . The cQMC energies (circles and squares) deviate from the underlying smooth trend (lines) by a quasi-random amount (of amplitude represented by the width of the shaded area around the lines). The smooth trend can be represented by a quadratic function, $E(w)$ (dashed line), for $w_{\text{h.o.}} < w < w_{\text{exh.}}$ (shaded middle region), while for $w < w_{\text{h.o.}}$ higher-order contributions become important, and for $w > w_{\text{exh.}}$ orbital-basis exhaustion sets in. At the value of w_0 corresponding to the exact wave function the quadratic function gives the exact energy, $E_0 = E(w_0)$.

We estimate the value of α by performing a preliminary least-squares fit to the bare data, $\tilde{E}(w)$, and evaluating

$$\alpha^2 = \max \left(0, \left\langle \left[E_i - \tilde{E}(w_i) \right]^2 \right\rangle - \langle (\Delta E_i)^2 \rangle \right), \quad (4.7)$$

i.e., we obtain α as the root-mean-square deviation of the data from the fit value not accounted for by statistical uncertainty alone. For this procedure to produce a meaningful result, the number of data points in each curve must be significantly greater than the

number of parameters in the quadratic fit function. We use at least 7 data points for all fits reported here.

In order to account for the statistical uncertainty and quasi-random fluctuations in the XSPOT method, we use a Monte Carlo resampling technique in which we generate 100000 instances of the VMC and DMC datasets in which a random amount proportional to $\sqrt{\alpha^2 + \Delta E_i^2}$ is added to the original energy values. We then perform fits to the data and find the intersection point for each such instance, and obtain the final result by averaging over instances. Figure 4.3 illustrates this process.

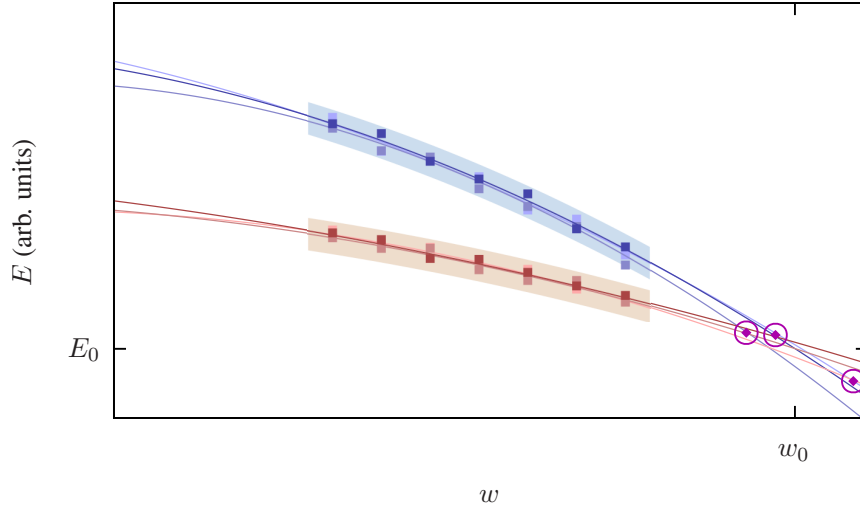


Figure 4.3: Illustration of the Monte Carlo resampling scheme used to compute statistics on the intersection between two curves. For each curve, having obtained an estimate of α (width of the shaded region) from the original energy data (not shown), we create a synthetic instance of the dataset by shifting the original points by a random amount proportional to $\sqrt{\alpha^2 + \Delta E_i^2}$ (squares of same color saturation), perform a quadratic fit (line), and find the intersection between both fits (circled diamond). This process is repeated over the random instances (three shown in the illustration), from which statistics on the intersection are obtained.

This procedure provides meaningful uncertainties on the intersection energies which account for both the cQMC statistical uncertainty and quasi-random deviations from the smooth trend. We demonstrate the full statistical procedure of the XSPOT method on multideterminant-Jastrow data for the carbon dimer in Figure 4.4. Notice that the distribution of intersection points shown in the inset of Figure 4.4 has a tail extending towards low E and large w . We find these tails to become more problematic the more parallel the

two intersecting curves are, eventually preventing the evaluation of an intersection point at all. It is therefore important to try to use curves which are as perpendicular as possible.

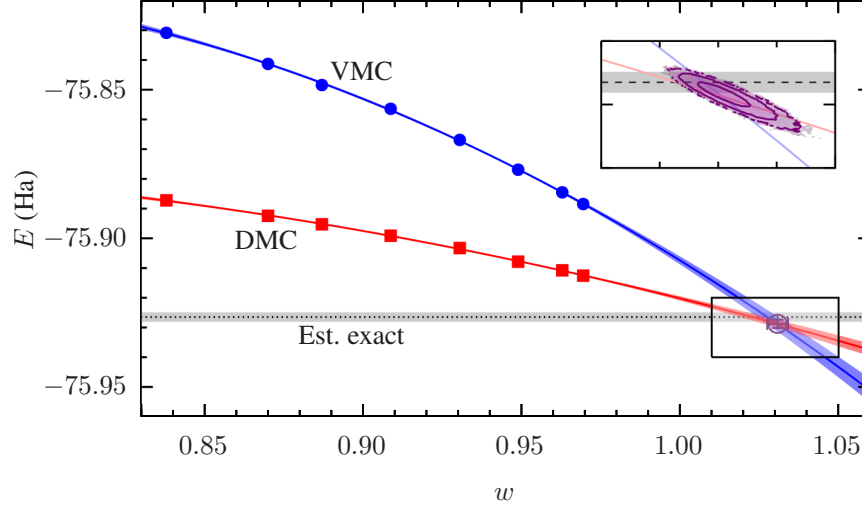


Figure 4.4: VMC and DMC energies of the ground-state C_2 molecule as a function of w , as shown in Figure 4.1, using the full statistical treatment of the XSPOT method. The translucent areas represent the uncertainty in the values of the fit at each point. Also shown is the estimated exact non-relativistic energy⁵ as a dotted line with a shaded area of ± 1 kcal/mol around it, and the intersection point between the curves. The inset shows the statistical distribution of intersection points as a color map with overlaid contour curves.

In Figure 4.5, we include additional VMC and DMC data using an inhomogeneous backflow transformation (“bVMC” and “bDMC”) for the carbon dimer. We also list the intersections between pairs of curves in Table 4.3, all of which are within uncertainty of each other. The VMC and backflow DMC curves provide the best-resolved results, which is to be expected since these curves intersect at the widest angle among all pairs of curves, as can be seen in Figure 4.5. By contrast, the DMC and bVMC curves incur a small but non-zero fraction of “missed” intersections, i.e., random instances of the data whose fits fail to intersect at $w > 1$, which signals the presence of heavy tails in the intersection distribution, resulting in a large uncertainty on the intersection energy.

For the four curves in Figure 4.5, the estimated amplitude of the quasi-random fluctuations ranges from $\alpha = 0.17$ to 0.41 mHa. In this work, we converge the cQMC calculations so that the uncertainties satisfy $\langle (\Delta E_i)^2 \rangle < \alpha^2$, i.e., so that quasi-random fluctuations represent the main contribution to the uncertainty on the fits and intersections. The statistical

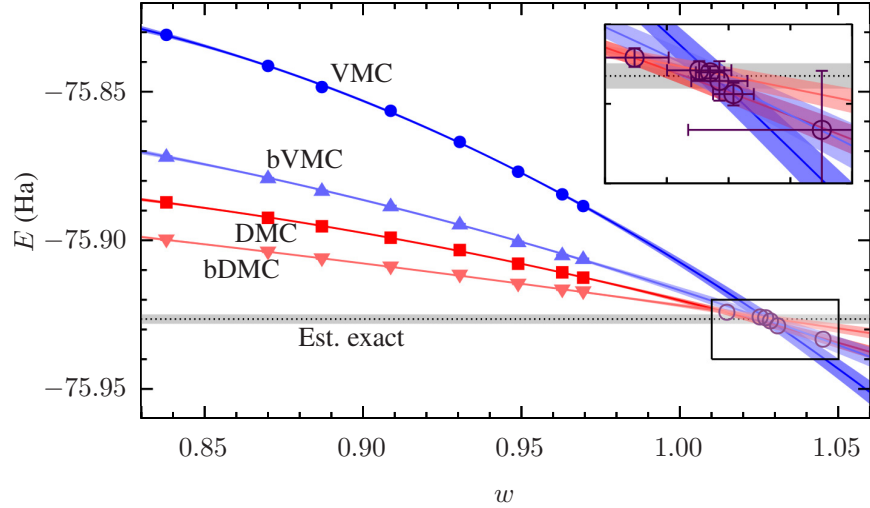


Figure 4.5: VMC and DMC energies of the ground-state C_2 molecule as a function of w , without (“VMC” and “DMC”) and with (“bVMC” and “bDMC”) a backflow transformation. The translucent areas represent the uncertainty in the values of the fit at each point. Also shown is the estimated exact non-relativistic energy⁵ as a dotted line with a shaded area of ± 1 kcal/mol around it, and the intersection points between each of the six possible pairs of curves; error bars on these are only shown in the inset for clarity.

Curves	w_0	E_0 (a.u.)	Miss (%)
VMC-DMC	1.031(3)	-75.9288 ± 0.0014	0.00
VMC-bVMC	1.028(5)	-75.9271 ± 0.0025	0.00
VMC-bDMC	1.027(2)	-75.9260 ± 0.0008	0.00
DMC-bVMC	1.05(2)	-75.9333 ± 0.0074	2.75
DMC-bDMC	1.015(6)	-75.9242 ± 0.0012	0.00
bVMC-bDMC	1.025(5)	-75.9258 ± 0.0012	0.00

Table 4.3: Location of all six pairwise intersections of the VMC, DMC, bVMC, and bDMC curves shown in Figure 4.5 for the C_2 molecule. “Missed intersections” refer to random instances of the curves that do not intersect at $w > 1$ in the Monte Carlo resampling procedure.

uncertainties on the cQMC energies can thus be neglected for all practical purposes.

4.2.3 Choice of Orbital Basis

The choice of orbitals plays a crucial role in the behavior of the cQMC energies as a function of w , modifying the point $w_{\text{exh.}}$ at which the effects of exhaustion start to become noticeable. In Figure 4.6, we demonstrate this for the H_2O molecule by comparing the cQMC energies obtained using Hartree-Fock orbitals expanded in the cc-pCVTZ basis set and natural orbitals expanded in the same basis set constructed so as to diagonalize the one-body density matrix in coupled cluster singles and doubles (CCSD).

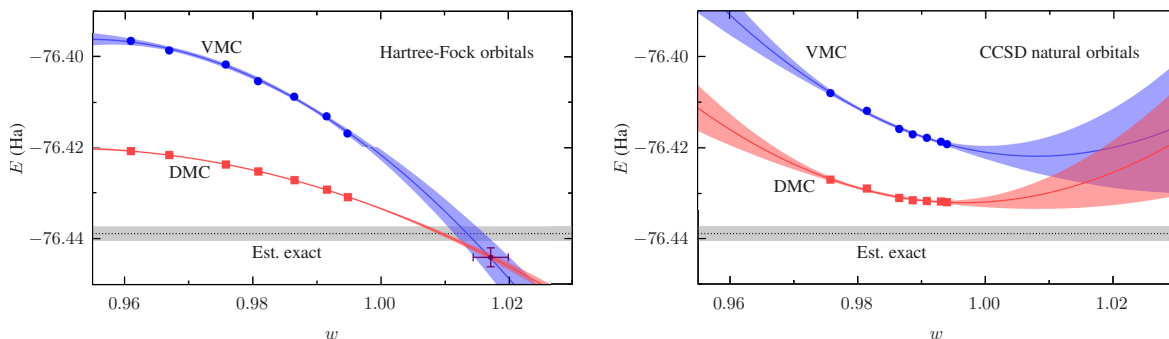


Figure 4.6: VMC and DMC energies of the H_2O molecule as a function of w , both using Hartree-Fock orbitals (left) and CCSD natural orbitals (right) expanded in the cc-pCVTZ basis set. In both cases the same numbers of CSFs are used. The translucent areas represent the uncertainty in the values of the fit at each point. Also shown is the estimated exact non-relativistic energy¹ as a dotted line with a shaded area of ± 1 kcal/mol around it, and the intersection point between the VMC and DMC curves in the left panel. The cQMC energies obtained with natural orbitals plateau with w , preventing the quadratic extrapolations from reaching the exact energy.

While the CCSD natural orbitals produce lower cQMC energies and larger values of w at fixed expansion sizes, the cQMC energies we obtain using Hartree-Fock orbitals follow the quadratic trend throughout the whole w range considered, while those obtained with natural orbitals plateau very early on, preventing their meaningful extrapolation.

4.3 Results and Discussion

In our calculations, we use Hartree-Fock orbitals expanded in the cc-pCVTZ Gaussian basis set^{40,41} obtained using MOLPRO.^{44,45} We perform a small-scale FCIQMC calculation

using the NECI package^{12,80} with configuration state functions (CSFs) instead of Slater determinants as walker sites,¹³⁰ which reduces the number of FCIQMC walkers required to accurately represent the wave function; note that the use of CSFs is not a requirement of the XSPOT method. The FCIQMC population is grown to 10^6 walkers and equilibrated, and the coefficients $\{c_I\}_{I=1}^{M_{\text{gen}}}$ of the $M_{\text{gen}} = 10\,000$ most-occupied CSFs are recorded from a population snapshot. From this information, we build CI expansions with the M CSFs whose coefficients are the largest in magnitude, where $M \leq 1500 \ll M_{\text{gen}}$.

The CSF coefficients are reoptimized in the presence of a Jastrow factor of the Drummond-Towler-Needs form,^{15,16} and of an optional inhomogeneous backflow transformation including electron-electron, electron-nucleus and electron-electron-nucleus terms.¹⁷ Note that even though CSFs are used, the presence of the Jastrow factor and of the backflow transformation prevents the cQMC trial wave function from formally being an exact spin state.⁷⁷ We optimize our wave function parameters using linear least-squares energy minimization^{110,111} with 10^6 statistically independent VMC-generated electronic configurations, a number large enough that the optimized cQMC energy can be assumed to lie reasonably close to its variational minimum; note that any remaining optimization error is absorbed into the quasi-random error q . The resulting trial wave function is then used to run two DMC calculations with time steps 0.001 and 0.004 a.u. and target populations of 2048 and 512 configurations, respectively, except for bDMC runs on CO_2 at 500 and 1000 CSFs for which we use 65536 and 16384 configurations. These energies are then linearly extrapolated to the zero time-step, infinite-population limit.^{14,70} We use the CASINO package^{14,72} to run the cQMC calculations, and use multi-determinant compression¹²⁸ to reduce the computational expense of evaluating the trial wave function.

We test the XSPOT method on all the eight systems under consideration to assess the different aspects discussed in Section 4.2 and to determine the broader applicability of the method. The cQMC energies and fits we obtain for the eight systems are shown in Figure 4.7. Also, the fraction of orbitals used, a proxy for the degree of orbital-basis exhaustion, is plotted for each of the systems in Figure 4.8.

We do not use all of the orbitals in the basis in any of our calculations, and the curves in Figure 4.7 do not seem to exhibit symptoms of orbital-basis exhaustion, except for the bVMC curves. We hypothesize that the backflow transformation makes some of the information contained in the orbitals redundant, which effectively reduces the value of $w_{\text{exh.}}$ with respect to the non-backflow data. Judging by the quality of the intersections, this issue affects bVMC more than it does bDMC, which would imply that this redundancy does not involve the location of the nodes of the trial wave function to the same degree as its values away from the nodes. This is in any case an empirical explanation; we do not use the bVMC data to obtain our final results.

In Table 4.4, we compare the total energies obtained from applying the xSPOT method to VMC and bDMC energy data with benchmark-quality estimates of their exact nonrelativistic values from the literature, along with our best backflow DMC result and prior cQMC results for reference. The corresponding relative energies are tabulated in Table 4.5.

System	Prior cQMC	Our best bDMC	xSPOT	Benchmark
C	$-37.84438(5)^a$	$-37.8442(0)$	-37.8439 ± 0.0003	-37.8450^b
N	$-54.58829(7)^a$	$-54.5881(0)$	-54.5897 ± 0.0014	-54.5893^b
O	$-75.06591(8)^a$	$-75.0647(1)$	-75.0664 ± 0.0004	-75.0674^b
H ₂ O	$-76.4389(1)^d$	$-76.4336(1)$	-76.4410 ± 0.0013	-76.4389^c
C ₂	$-75.9229(6)^a$	$-75.9172(1)$	-75.9260 ± 0.0008	-75.9265^b
C ₂ *		$-75.8374(1)$	-75.8455 ± 0.0011	$-75.8465^{e,b}$
N ₂	$-109.5372(3)^a$	$-109.5344(1)$	-109.5394 ± 0.0018	-109.5425^c
CO ₂		$-188.5837(2)$	-188.5983 ± 0.0043	-188.6015^c

^aRef. 116; ^bRef. 5; ^cRef. 1; ^dRef. 131; ^eRefs. 132–134.

Table 4.4: Total energies in Ha, obtained with the xSPOT method using the VMC and bDMC data for the various atoms and molecules considered in this work, along with results from prior multi-determinant cQMC studies, our best individual bDMC energy for each system, and benchmark-quality non-relativistic total energies from the literature.

	xSPOT	Benchmark
H ₂ O $\rightarrow$ 2H + O	374.6 ± 1.4	371.4^b
C ₂ $\rightarrow$ 2C	238.3 ± 0.9	236.5^a
C ₂ $\rightarrow$ C ₂ *	80.6 ± 1.3	80.0^c
N ₂ $\rightarrow$ 2N	360.1 ± 2.3	363.9^b
CO ₂ $\rightarrow$ C + 2O	621.6 ± 4.3	621.7^b

^aRef. 5; ^bRef. 1; ^cRefs. 132–134.

Table 4.5: Atomization and excitation energies in mHa of the various molecules considered in this work corresponding to the total energies in Table 4.4 obtained from the xSPOT method, along with benchmark-quality non-relativistic relative energies from the literature.

For the excited-state C₂*, we instead compare the vertical excitation energy with that calculated with internally-contracted multi-reference coupled cluster theory (ic-MRCC);^{132–134}

we compute the total energy of C_2^* shown in Table 4.4 by adding the ic-MRCC excitation energy to the estimated ground-state energy of C_2 from Ref. 5. All of the total energies reported in Table 4.4 are within statistical uncertainty of their corresponding benchmark values. An important observation is that our individual cQMC energies are not lower than those from prior cQMC calculations, implying that we incur a lower computational cost, but our XSPOT results are in general closer to the benchmarks than cQMC results from prior studies. Our XSPOT energies are on average 1.1 standard error above the benchmark, with a root-mean-square deviation of 1.9 standard errors. These results are compatible with the XSPOT method being exact when the method’s assumptions are satisfied. The relative energies are likewise in agreement with the benchmark values.

We find that the magnitude α of the quasi-random fluctuations of the cQMC energies is of up to 0.7 mHa. These fluctuations are particularly visible in the VMC data for N, O, and H_2O in Figure 4.7, for example. The magnitude of the quasi-random fluctuations does not seem to increase too rapidly with system size, but their effect on the extrapolated energy becomes more pronounced the further the cQMC data are from the intersection. This increasing uncertainty on the XSPOT total energies, reaching 4 mHa for the CO_2 molecule, hints at a limitation of the XSPOT method. The cQMC energies and values of w obtained using modest-sized multi-determinant expansions with a fixed basis set move away from the intersection point with increasing system size, which in turn amplifies the effects of quasi-random noise on the uncertainty of the XSPOT energy. One would have to use bigger basis sets and larger multi-determinantal expansions to get data closer to the intersection in order to reduce this uncertainty, increasing the computational cost of the method.

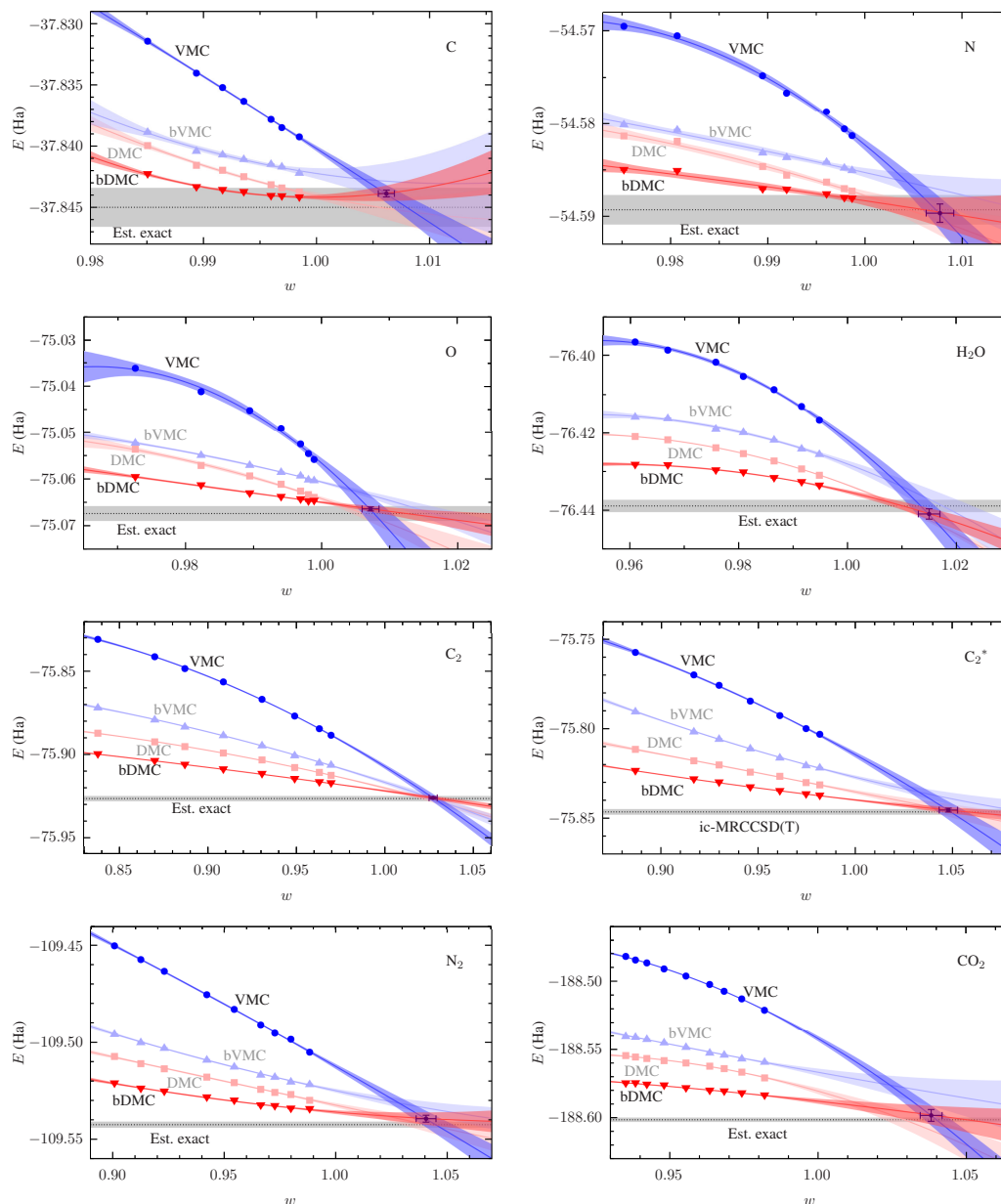


Figure 4.7: VMC and DMC energies of the atoms and molecules considered in this work as a function of w , without (“VMC” and “DMC”) and with (“bVMC” and “bDMC”) a backflow transformation. The translucent areas represent the uncertainty in the values of the fit at each point. Also shown in each plot is the relevant benchmark energy (see details in text and Table 4.4) as a dotted line with a shaded area of ± 1 kcal/mol around it, and the intersection point between the VMC and bDMC curves.

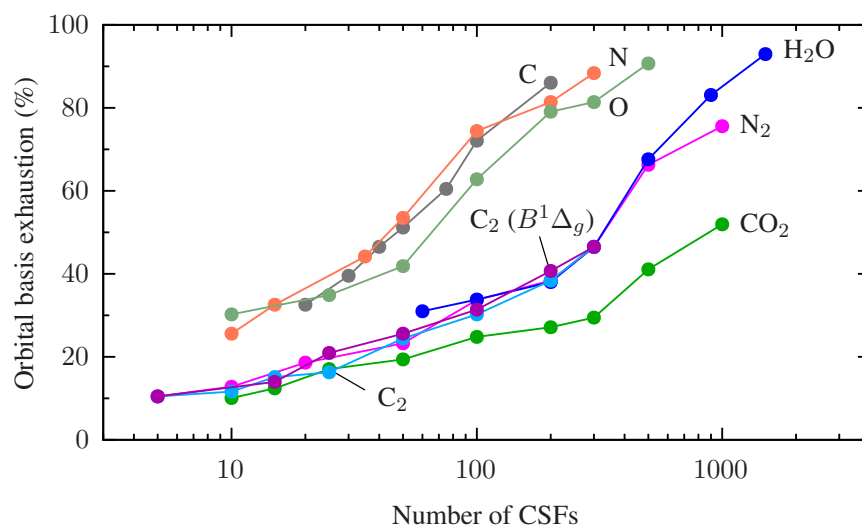


Figure 4.8: Number of distinct spatial orbitals in the FCIQMC trial wave functions relative to that in the orbital basis as a function of the number of CSFs, for the various systems considered in this work.

Chapter 5

Correlation Energy of the High-density Spin-unpolarized Electron Gas at the Thermodynamic Limit

We employ methodology recently proposed by Ruggeri *et al.*²¹ that combines the DMC and FCIQMC methods, along with a finite-size extrapolation framework, to calculate the thermodynamic limit of the exact correlation energy of uniform electron gases, which we apply to the high-density spin-unpolarized case. Using the accurate transcorrelated FCIQMC (TC-FCIQMC) method, the thermodynamic limit is obtained as -77.74(54) mHa at $r_s = 0.5$, which is within uncertainty of the recently reported correlation energy from the literature. The fixed-node error is also extrapolated and found to exceed 4 mHa.ⁱ

5.1 Introduction

The three-dimensional uniform electron gas (UEG) plays a crucial role in the understanding of the nature of electronic correlations in real materials²⁰. The correlation energy of the 3D UEG as a function of its density is a fundamental element in the description of the behavior of electrons in real systems by density functional theory (DFT)^{6,7}.

ⁱThis chapter is based on a collaborative work with Pablo López Ríos and Ali Alavi.

Although it appears simple at first glance, the intricate correlations arising from the long-range nature of the Coulomb interaction necessitate the use of many-body wave-function-based methods, such as quantum Monte Carlo (QMC) techniques^{11,17,21,95,135–141}, for an accurate characterization of the UEG. The release-node DMC simulations of Ceperley and Alder (CA)¹¹ provided important data connecting the high- and low-density regimes of the correlation energy. Their data enabled the development of parametrizations of the correlation energy over a wide density range which are frequently used in DFT calculations. Prominent parameterizations that make use of Ceperley and Alder’s results were carried out by Vosko, Wilk, and Nusair (VWN80)¹⁴², Perdew and Zunger (PZ81)¹⁴³, and Perdew and Wang (PW92)¹⁴⁴. The PW92 parametrization, the most popular functional form in DFT calculations, includes five parameters, two of which are determined from analytic high-density constraints and three by fitting to the CA data. Later, a “density parameter interpolation” (DPI)¹⁴⁵ parametrization was introduced that is constructed by imposing four high-density and three low-density constraints on a seven-parameter functional form.

In the present work, we provide new QMC data for the correlation energy of the high-density spin-unpolarized three-dimensional UEG. We use a combination of the transcorrelated full configuration interaction quantum Monte Carlo (TC-FCIQMC) and fixed-node diffusion Monte Carlo (DMC) methods to compute the correlation energy of the spin-unpolarized three-dimensional UEG at $r_s = 0.5$. We employ a recently proposed extrapolation method which is found to be much more accurate than previous approaches²¹. By extrapolating the fixed-node energy and the fixed-node error to the thermodynamic limit, we obtain the correlation energy at $r_s = 0.5$, which is within uncertainty of the QMC correlation energy recently reported at this density¹³⁸.

The variational Monte Carlo (VMC)^{10,14} and fixed-node DMC^{11,14} methods have been extensively used to study the UEG^{17,21,137–141}. In the VMC method, parameters in an approximate trial wave function are optimized according to the variational principle with energy expectation values computed by Monte Carlo integration in real space. In the DMC method, the imaginary-time Schrödinger equation is used to evolve a statistical ensemble of electronic configurations in real space towards the ground state. The antisymmetry principle is satisfied using the fixed-node approximation in which the intricate nodal surface of the wave function is constrained to equal that of an approximate trial wave function optimized in VMC. This incurs a positive bias in the calculated energy, referred to as the fixed-node error.

The FCIQMC^{12,13} method is a second-quantized projection technique which has also been used for the accurate study of the UEG^{21,95,135,136}. The FCIQMC method explicitly operates in the basis of Slater determinants where random walkers sample the discrete

Hilbert space. The TC-FCIQMC method⁹⁵ is the incorporation of transcorrelation into the FCIQMC method. The transcorrelated (TC) method is based on a similarity transformation of the Hamiltonian, offering substantial improvements in basis-set convergence and in the effective level of theory of the underlying correlation method^{95–106}.

The remaining sections of this chapter are organized as follows. In the next section, we describe the theoretical details and the extrapolation technique, which is based on the methodological section of Ref. 21. Then, we present our calculations on the spin-unpolarized UEG systems and we discuss the results in Section 5.3.

5.2 Methodology

The second-quantized Hamiltonian of an infinite UEG is written as

$$\hat{H} = \frac{1}{2} \sum_{\mathbf{k}} k^2 \hat{a}_{\mathbf{k}}^\dagger \hat{a}_{\mathbf{k}} + \sum_{\mathbf{p}, \mathbf{q}} \sum_{\mathbf{k} \neq 0} \frac{4\pi}{k^2} \hat{a}_{\mathbf{p}+\mathbf{k}}^\dagger \hat{a}_{\mathbf{q}-\mathbf{k}}^\dagger \hat{a}_{\mathbf{q}} \hat{a}_{\mathbf{p}}, \quad (5.1)$$

where $\mathbf{k}$, $\mathbf{p}$, and $\mathbf{q}$ refer to reciprocal-space vectors, and $\hat{a}_{\mathbf{k}}^\dagger$ and $\hat{a}_{\mathbf{k}}$ denote the creation and annihilation operators for a single-electron state with a wave vector of $\mathbf{k}$, respectively. The UEG is characterized by its uniform number density n , defined via $r_s = (4\pi n/3)^{-1/3}$. The kinetic energy term is diagonal, and the interaction term only establishes connections between states with equal total momentum $\mathbf{k}_T = \sum_j \mathbf{G}_{\mu_{Ij}}$, where $\mathbf{G}$ is the reciprocal lattice vector and μ_{Ij} is the index of the j th wave vector occupied in the I th determinant ($I = 1$ corresponds to the HF determinant). The Hilbert space of the UEG then consists of disjoint subspaces corresponding to different amounts of $\mathbf{k}_T$, and the ground state is the solution to the Schrödinger equation within the subspace where the total energy is minimized.

5.2.1 Twist Averaging and Brillouin Zone Division

The finite-size error in the energy of the UEG comprises a smoothly varying contribution with the number of electrons N and a quasi-random component that needs to be removed for a clear extrapolation of the smooth part²¹. Twist averaging^{33–35} is a technique which significantly diminishes quasi-random fluctuations by averaging over wave-vector offsets $\mathbf{k}_s$ in the Brillouin zone. Changes in the twist $\mathbf{k}_s$ result in discrete variations for the total momentum of the ground-state wave function $\mathbf{k}_T = \sum_j \mathbf{G}_{\mu_{1j}}$, dividing the Brillouin zone into Z regions associated with different amounts of $\mathbf{k}_T$ ²¹.

If for a fixed $\mathbf{k}_T$ the reciprocal lattice is shifted by $\mathbf{k}_s$, according to the Equation 5.1, a constant is added to the diagonal kinetic energy term, and the interaction term is left unchanged as it only depends on differences between reciprocal lattice vectors. The correlation energy thus only relies on the total momentum $\mathbf{k}_T$, which varies discretely with $\mathbf{k}_s$. Therefore, the twist-averaged correlation energy can be written as

$$\begin{aligned} E_{\text{corr}}^{\text{TA}} &= \frac{1}{\Omega_{\text{BZ}}} \int_{\text{BZ}} E_{\text{corr}}(\mathbf{k}_s) d\mathbf{k}_s \\ &= \sum_z \frac{\Omega_z}{\Omega_{\text{BZ}}} E_{\text{corr}}(\mathbf{k}_s^z), \end{aligned} \quad (5.2)$$

where Ω_{BZ} and Ω_z are the volume of the Brillouin zone and the volume of the z th region of the Brillouin zone, respectively, and $\mathbf{k}_s^z$ is an arbitrary twist shift in the z th region. This method requires dividing the Brillouin zone and calculating parameters associated with each of the Z regions within the Brillouin zone. The division can be performed by minimizing the energy of the non-interacting UEG system³³.

The energy (per electron) of the non-interacting electron gas is equal to the HF kinetic energy, and can be written as

$$\begin{aligned} E_{\text{NI}}(\mathbf{k}_s, \mathbf{k}_T) &= K_1(\mathbf{k}_s, \mathbf{k}_T) \\ &= \frac{1}{2N} \sum_{j=1}^N (\mathbf{G}_{\mu_{1j}} + \mathbf{k}_s)^2 \\ &= \frac{1}{2N} \sum_{j=1}^N (|\mathbf{G}_{\mu_{1j}}|^2 + |\mathbf{k}_s|^2 + 2\mathbf{G}_{\mu_{1j}} \cdot \mathbf{k}_s) \\ &= E_{\text{NI}}(\mathbf{0}, \mathbf{k}_T) + \frac{1}{2} |\mathbf{k}_s|^2 + \frac{1}{N} \mathbf{k}_T \cdot \mathbf{k}_s, \end{aligned} \quad (5.3)$$

where μ_{1j} is the index of the j th wave vector occupied in the HF determinant. The energy of this system at fixed $\mathbf{k}_T$ forms a paraboloid centered around $\mathbf{k}_s = -\frac{1}{N}\mathbf{k}_T$. The total momentum $\mathbf{k}_T$ changes discretely at the intersection of two such paraboloids, such that if we label the total momenta of two adjacent regions as $\mathbf{k}_T^{z_1}$ and $\mathbf{k}_T^{z_2}$, the intersections gives

$$\frac{1}{N} (\mathbf{k}_T^{z_1} - \mathbf{k}_T^{z_2}) \cdot \mathbf{k}_s = E_{\text{NI}}(\mathbf{0}, \mathbf{k}_T^{z_2}) - E_{\text{NI}}(\mathbf{0}, \mathbf{k}_T^{z_1}). \quad (5.4)$$

This represents the equation of a plane. Hence, the regions of constant total momentum within the Brillouin zone (the irreducible Brillouin zone (IBZ) is used in practice) are

convex polyhedra. This equation characterizes a Bragg plane, and the intersections of inter-region and/or IBZ planes give the region vertices.

For a given shift $\mathbf{k}_s$, identifying the N reciprocal lattice vectors with the smallest $|\mathbf{G}_i + \mathbf{k}_s|$ provides the indices of the occupied orbitals $\{\mu_{1j}\}$, subsequently determining $\mathbf{k}_T$. However, there is not a unique set of occupied orbitals at points located on inter-region planes, and multiple total momenta result in the same kinetic energy. Then, the allowed values of the total momentum at a vertex can be determined by considering all possible occupations. Each pair of planes intersect along a line corresponding to a polyhedral edge, indicating an adjoining vertex. Consequently, it is feasible to identify the vertices of all polyhedral regions in the IBZ by progressively transitioning between neighboring vertices along the region edges. The exact volume and center of the IBZ regions can be further obtained using the divergence theorem²¹. After performing the exact division, we run TC-FCIQMC and DMC calculations at each of the IBZ regions.

The requirement to enumerate all possible occupations of partially-filled shells poses a computational bottleneck in the exact division algorithm, limiting its practical application to cases with $N \lesssim 100$ ²¹. For larger UEG systems, we sample the Brillouin zone randomly in order to calculate the twist-averaged DMC correlation energy.

5.2.2 Basis-set Extrapolation of TC-FCIQMC Energies

In the high-density regime, where the UEG behaves as a Fermi liquid, plane waves with the smallest wave vectors are regarded as good choices for the basis sets. These finite basis sets can be extrapolated to the complete-basis-set (CBS) limit⁹. Given that, we make use of the following fitting function²¹ to extrapolate the TC-FCIQMC correlation energies at each of the IBZ regions z .

$$E_{\text{corr}}^z(M) = E_{\text{corr}}^z(\infty) + a_z M^{-1} + b_z M^{-2}, \quad (5.5)$$

where M is the basis-set size, and $E_{\text{corr}}^z(\infty)$, a_z , and b_z are fit parameters. Since the basis-set error is roughly independent of z ²¹, we perform extrapolation by setting the parameters $a_z = a$ and $b_z = b$ across all Z regions. Figure 5.1, as an example, illustrates the finite basis-set error in the TC-FCIQMC correlation energy as a function of M^{-1} for the unpolarized 38-UEG at $r_s = 0.5$. Furthermore, when the extrapolation is done, we perform twist averaging via Equation 5.2.

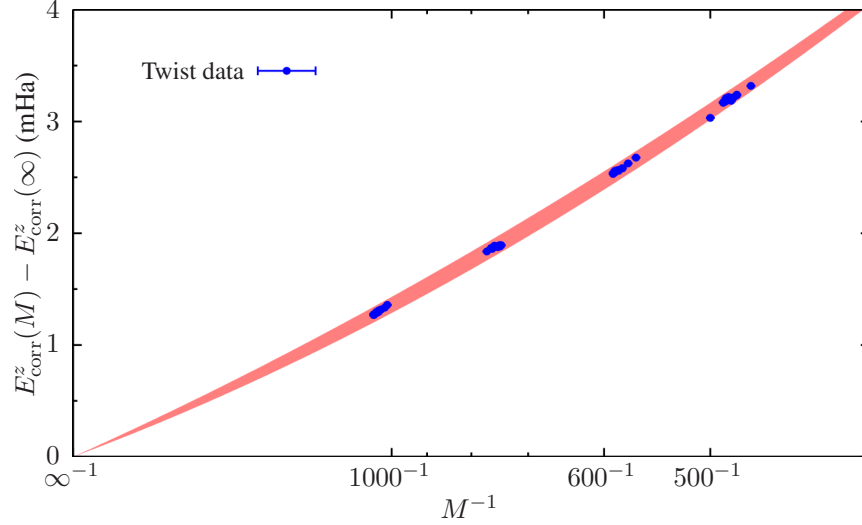


Figure 5.1: Finite basis-set error in the TC-FCIQMC correlation energy as a function of inverse basis-set size M^{-1} for the unpolarized 38-UEG at $r_s = 0.5$. The data correspond to calculations at multiple basis-set sizes in each of the IBZ regions. The translucent areas represent the uncertainty in the values of the fit at each point.

5.2.3 Finite-size Extrapolation Framework for Fixed-node DMC

Apart from twist averaging, other sources of finite-size errors in DMC energies need to be accounted for. Two significant components stem from the HF exchange and total DMC energies. The leading-order contribution to the finite-size error in the HF exchange energy is equal to $hN^{-2/3}$, where $h = -\frac{3\epsilon_1\sqrt[3]{2}}{16\pi}r_s^{-1}$ for the unpolarized UEG and $\epsilon_1 = 5.674594959$ for a simple cubic simulation cell^{21,35}. Besides, the leading-order contribution to the finite-size error in the total DMC energy equals tN^{-1} , where $t = -\frac{\sqrt{3}}{2}r_s^{-3/2}$ for the unpolarized UEG³⁴. It is found²¹ that the smooth part of the finite-size error can be represented as a polynomial in $N^{-1/3}$, and as the beyond-leading-order contributions to the finite-size error in the HF kinetic and exchange energies are proportional to $N^{-4/3}$, the DMC correlation energy can be extrapolated as^{21,35}

$$E_{\text{corr}}^{\text{FN}}(N) + hN^{-2/3} - tN^{-1} = c_0 + c_1N^{-4/3} + c_2N^{-5/3} + c_3N^{-2} + \dots, \quad (5.6)$$

where $\{c_n\}$ are density-dependent parameters. This equation is used to estimate the thermodynamic limit of the fixed-node DMC correlation energy. Moreover, it is shown¹⁴⁶ that the contribution of the backflow functions¹⁷ to the finite-size error in the UEG energy is at the order of N^{-1} . Also, one can consider that the fixed-node error exhibits a similar

asymptotic behavior as the backflow contribution to the energy²¹, aligning with the observed approximate proportionality between these two quantities^{114,139}. With expecting that the fixed-node error exhibits a weaker variation with N compared to the fixed-node energy, we can model it using a lower-order expression as²¹

$$\varepsilon_{\text{FN}}(N) = f_0 + f_1 N^{-1} + f_2 N^{-4/3} + \dots, \quad (5.7)$$

where ε_{FN} is the fixed-node error and $\{f_n\}$ are density-dependent parameters.

5.3 Results and Discussion

We simulate finite systems consisting of N electrons, with an equal number of spin-up and spin-down electrons, in a simple cubic simulation cell of side $L = (N/n)^{1/3}$ at fixed homogeneous densities. In the VMC calculations, we use the Slater-Jastrow^{15,16} form of the trial wave function. We optimize the Jastrow factor parameters by minimizing the VMC energy^{110,111} and minimizing the mean absolute deviation of local energies from the median¹⁴⁷. The optimized Jastrow factor is used to run two DMC calculations with time steps of $0.01r_s^2$ and $0.04r_s^2$ a.u. and target populations of 2048 and 512 configurations, respectively. The obtained energies are then linearly extrapolated to the zero time-step limit^{14,70}. The VMC and DMC calculations are performed using the CASINO^{14,72} package, and we report energies per electron.

To twist average DMC correlation energies, at system sizes $14 \leq N \leq 66$ we perform region-based twist averaging by running DMC calculations in each of the IBZ regions, while for $114 \leq N \leq 778$ we perform Monte-Carlo twist averaging using 32 random values of $\mathbf{k}_s$. Table 5.1 tabulates the obtained twist-averaged DMC correlation energies at $r_s = 0.5$, 1.0, and 2.0 with corresponding thermodynamic limits, which are estimated via Equation 5.6 by setting h and t to their analytic values and considering c_0 , c_1 , c_2 , and c_3 as fit parameters. In Figure 5.2, we plot $E_{\text{corr}}^{\text{FN}}(N) + hN^{-2/3} - tN^{-1}$ at $r_s = 0.5$, 1.0, and 2.0 as a function of N^{-1} , along with corresponding extrapolation fits to the thermodynamic limits. These extrapolations numerically validate that there are no additional contributions to the Equation 5.6 at the order of N^{-1} or slower²¹.

Our transcorrelated FCIQMC calculations are performed with up to two-body terms (refer to Equations 3.60 and 3.61), using RPA approximations to treat the three-body terms⁹⁵. We also make use of the initiator approximation¹³, which allows a substantial reduction in the number of walkers n_w required for the convergence of the TC-FCIQMC energy. However, this advantage introduces a systematic bias, known as the initiator

N	$E_{\text{corr}}^{\text{FN}}$		
	$r_s = 0.5$	$r_s = 1.0$	$r_s = 2.0$
14	-32.451(4)	-29.750(5)	
30	-40.905(5)	-36.585(7)	
38	-43.904(3)	-38.860(4)	-32.278(4)
54	-47.637(3)	-41.642(3)	
66	-50.018(2)	-43.364(2)	-35.265(2)
114	-54.700(86)	-46.613(34)	-37.307(15)
162	-57.382(29)	-48.330(19)	-38.366(11)
186	-58.540(21)	-49.065(21)	
246	-60.419(34)	-50.304(9)	-39.515(4)
294	-61.462(28)	-50.977(19)	-39.881(11)
406	-63.326(14)	-52.096(12)	-40.532(4)
502	-64.438(16)	-52.719(22)	
610	-65.332(13)	-53.292(4)	
778	-66.409(10)	-53.914(7)	
∞	-73.681(18)	-57.913(11)	-43.693(16)

Table 5.1: Twist-averaged DMC correlation energies of the unpolarized UEG at $r_s = 0.5$, 1.0, and 2.0 and their respective thermodynamic limits, in mHa.

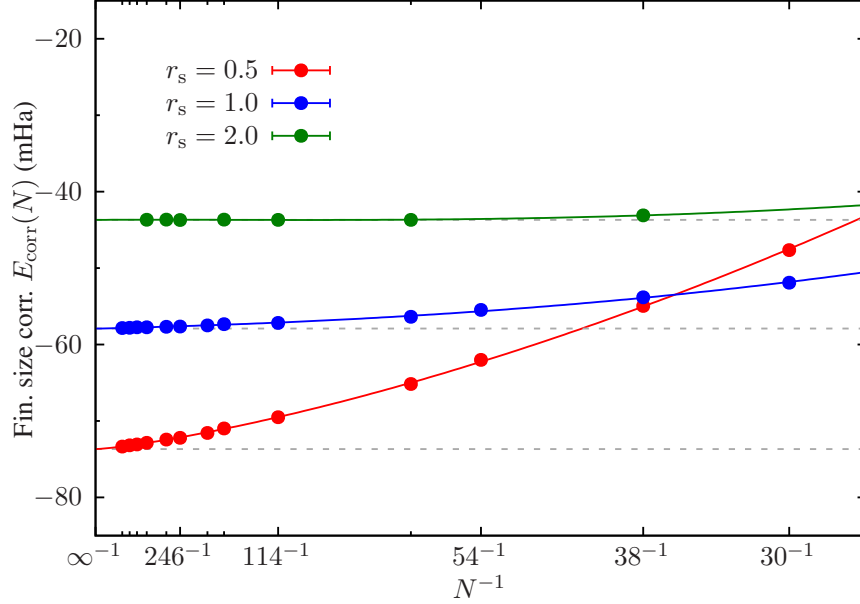


Figure 5.2: Finite-size corrected fixed-node correlation energies of the unpolarized UEG at $r_s = 0.5$, 1.0, and 2.0 as a function of N^{-1} . Horizontal dashed lines are shown at the thermodynamic limits for each density. The uncertainty in the fitted curves is smaller than the line width.

error, which can be minimized by increasing the walker population, and in practice, we increase the walker population until changes in the energy become negligible. The walker convergence study suggests $n_w = 2 \times 10^7$, 4×10^7 , 5.2×10^8 , and 5.2×10^8 walkers as suitable number of walkers to overcome the initiator error for the 14-, 30-, 38-, and 54-UEG systems at $r_s = 0.5$, respectively. We perform the TC-FCIQMC calculations using the NECI package^{12,80}.

We estimate the exact correlation energy by twist averaging the TC-FCIQMC correlation energies as explained in Section 5.2, at system sizes $N = 14, 30, 38$, and 54 at $r_s = 0.5$. Then, we calculate the fixed-node errors as the difference between the twist-averaged fixed-node DMC and the twist-averaged TC-FCIQMC correlation energies. The thermodynamic limit of the fixed-node error is also obtained using Equation 5.7 by treating f_0 , f_1 , and f_2 as fit parameters. In Figure 5.3, we plot the fixed-node error as a function of N^{-1} , along with the resulting fit curve. The fixed-node errors can be seen to exhibit non-monotonic behavior as a function of inverse system size, failing to be a well-behaved, smooth function as found in Ref. 21. We believe that obtaining more converged TC-FCIQMC correlation energies using our newly proposed walker number extrapolation technique in the next chap-

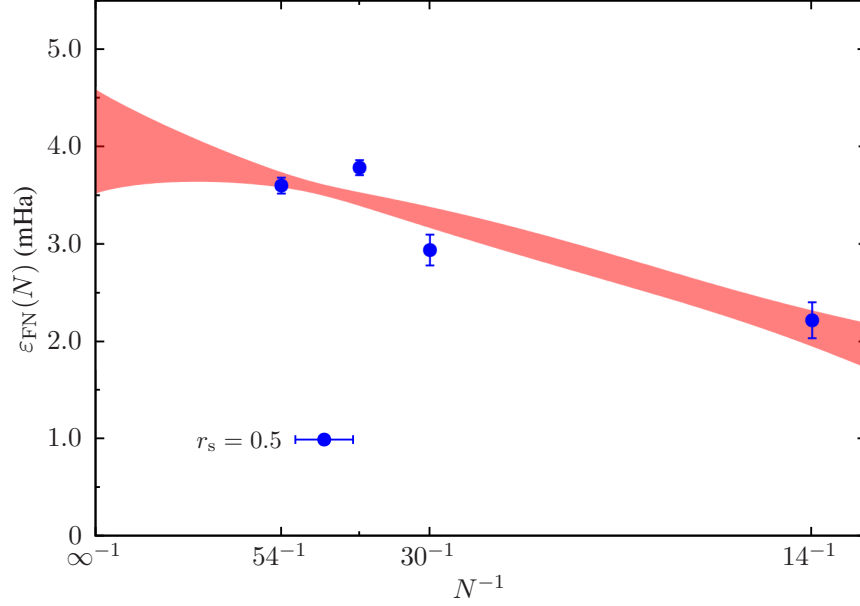


Figure 5.3: Fixed-node error for the unpolarized UEG at $r_s = 0.5$ as a function of N^{-1} . The linewidth represents the statistical uncertainty in the fit.

ter, might resolve this non-monotonic behavior, and reduce the statistical uncertainty of the thermodynamic limit.

Table 5.2 shows the obtained exact correlation energies and fixed-node errors at $r_s = 0.5$ with corresponding thermodynamic limits. We also compare our result at $r_s = 0.5$ with existing parametrizations and the most accurate DMC studies from the literature. As illustrated in Table 5.3, the correlation energy computed in this study, despite having a relatively large error bar, falls within the uncertainty range of the most recent correlation energy¹³⁸.

N	E_{corr}	ε_{FN}
14	-34.67(18)	2.22(18)
30	-43.84(16)	2.94(16)
38	-47.69(8)	3.78(8)
54	-51.24(8)	3.60(8)
∞	-77.74(54)	4.06(54)

Table 5.2: Twist-averaged TC-FCIQMC (“exact”) correlation energies and fixed-node errors for the unpolarized UEG at $r_s = 0.5$ for different system sizes and corresponding thermodynamic limits, in mHa.

	This work	VWN80 ¹⁴²	PZ81 ¹⁴³	PW92 ¹⁴⁴	DPI ¹⁴⁵	DMC ¹³⁷	DMC ¹³⁸
E_{corr}	-77.74(54)	-77.06	-76.03	-76.62	-77.45	-73.35	-78.06(7)

Table 5.3: Correlation energy of the unpolarized UEG at $r_s = 0.5$ obtained in this work, along with results from the VWN80, PZ81, PW92, and DPI parametrizations, and prior DMC studies, in mHa.

Chapter 6

Walker Number Extrapolation of FCIQMC Energies

A novel extrapolation scheme is introduced to eliminate the initiator error in the full configuration interaction quantum Monte Carlo (FCIQMC) method. We find that extrapolating the FCIQMC energy linearly in $n_w^{-1/3}$ to $n_w^{-1/3} \rightarrow 0$ removes the energy bias caused by the initiator approximation and converges the energy to the full configuration interaction (FCI) limit. We then apply this extrapolation technique to a series of first-row dimers including C_2 , CN , N_2 , and O_2 in the Dunning basis sets cc-pVnZ ($n=D, T, Q$, and 5), which yields results in excellent agreement with FCI limits estimated from the semistochastic heat-bath CI (SHCI) method. This method is also applied to investigate the more complex chromium atom and dimer in the cc-pVnZ-DK ($n=D, T$, and Q) basis sets.ⁱ

6.1 Introduction

The full configuration interaction quantum Monte Carlo (FCIQMC) algorithm¹² is a projective QMC method formulated to tackle the electronic Schrödinger equation within a given basis set. This method relies on the population dynamics of a set of walkers with positive and negative weights, and its master equation is obtained by interpreting the imaginary-time Schrödinger equation as a first-order kinetic equation. In the large-walker-

ⁱThis chapter is based on Refs. 98 and 148. The methodological work was done by Pablo López Ríos and Ali Alavi.

number limit, the population dynamics effectively samples the exact fermionic ground-state wave function.

While the FCIQMC algorithm in its full form has the capability to converge to the exact ground state, achieving this convergence necessitates a minimum number of walkers to ensure stable dynamics and to get the wave function’s sign structure to emerge¹⁸. This requisite minimum varies based on the specific system under examination and constitutes a fraction of the full Hilbert space size, yet it still scales proportionally, making the method unfeasible for larger systems.

The initiator approximation (i-FCIQMC)¹³ addresses this limitation by introducing a systematically improvable bias. In the initiator approximation, the walker dynamics are constrained, allowing only those walkers associated with determinants possessing sufficient population (referred to as initiators) to spawn new walkers onto unoccupied determinants. Spawns from non-initiators to unoccupied determinants are rejected. These constrained dynamics serve to prevent potential noise propagation and contamination of the wave function by non-initiators, ensuring stability in simulations regardless of the number of walkers. Consequently, i-FCIQMC becomes capable of handling significantly larger systems than the original FCIQMC algorithm. However, this advantage comes with a trade-off in the form of a systematic bias in the computed energy known as the initiator bias. This bias can be minimized to an arbitrary extent by increasing the number of walkers, and the algorithm is specifically designed to approach the FCI result in the infinite-walker limit. In practical terms, achieving convergence with respect to the walker number is feasible for systems containing up to about 20 electrons, well before memory requirements make the method impractical. As the system size expands, convergence with walker number slows down, reaching a point where it becomes practically impossible to converge to the exact FCI limit.¹⁴⁹

To ameliorate the initiator bias, several methods have been proposed so far. Ten-no¹⁵⁰ argues that most of the initiator error arises from a deficiency in proper scaling with system size that can be greatly mitigated by a size-consistency correction via stochastic coupled-electron pair approximations. The adaptive shift method^{149,151} reduces the initiator bias at a much lower number of walkers by giving each non-initiator determinant its own local shift, which is a proportionately scaled value of the overall global shift. Another technique to reduce the initiator error would be the use of the second-order¹⁵² or the third-order¹⁵³ Epstein–Nesbet perturbative correction to the initiator approximation.

To eliminate the initiator error, we present a new extrapolation method in which the FCIQMC total energy calculated at different walker numbers is extrapolated linearly in $n_w^{-1/3}$ to the infinite-walker limit. The application of this technique to a series of first-row

diatomics yields results matching those from the SHCI method²². The Cr atom and dimer are investigated, with results in agreement with previous studies. To this end, we first present the theoretical details of our method in Section 6.2. We then apply our method to the molecular systems, and we report and discuss the results in Section 6.3.

6.2 Methodology

The solution to the Schrödinger equation of an electronic system, $\hat{H}|\Phi\rangle = E|\Phi\rangle$, can be expressed as a full configuration interaction wave function,

$$|\Phi\rangle = \sum_{i=0}^{n_{\text{FCI}}} C_i |i\rangle, \quad (6.1)$$

where $\{|i\rangle\}$ are the set of $n_{\text{FCI}} + 1$ possible Slater determinants that can be constructed for the system in a given one-electron orbital basis, and $\{C_i\}$ are coefficients. To keep things simple, we shall assume the Hamiltonian to be Hermitian, $\hat{H} = \hat{H}^\dagger$, and that the reference energy is zero, $\langle 0|\hat{H}|0\rangle = 0$, throughout this chapter. The FCIQMC method stochastically samples the space of Slater determinants by assigning a signed number of walkers N_i to each determinant $|i\rangle$, and subjecting N_i to stochastic imaginary-time evolution according to the master equation

$$-\frac{dN_i}{dt} = (H_{ii} - S)N_i + \sum_{j \neq i} H_{ji}N_j, \quad (6.2)$$

where $H_{ji} = \langle j|\hat{H}|i\rangle$ and S is a shift parameter. In the infinite-walker limit, the walker numbers converge to $N_i \propto C_i$, the shift S tends to E , and the master equation trivially reduces to the time-independent Schrödinger equation,

$$0 = (H_{ii} - E)C_i + \sum_{j \neq i} H_{ji}C_j. \quad (6.3)$$

The initiator approximation modifies the dynamics imposed by the master equation to allow stabilizing calculations at a finite target number of walkers N , converging onto a truncated configuration interaction (CI) wave function containing n_{det} determinants,

$$|\tilde{\Phi}\rangle = \sum_{i=0}^{n_{\text{det}}} \tilde{C}_i |i\rangle, \quad (6.4)$$

and producing a measure of the energy of the system $\tilde{E}(N)$ which converges to E as $N \rightarrow \infty$. The purpose of this work is to find the leading-order behavior of the initiator bias $\tilde{E}(N) - E$ with N in the limit $N \rightarrow \infty$.

6.2.1 Modeling the Initiator Approximation

The initiator approximation is implemented by categorizing determinants as “initiators” when their population $|N_i|$ exceeds a predefined number of walkers N_{init} (typically about 3), and preventing non-initiators from spawning onto unpopulated determinants, i.e., those with $N_i = 0$. Walker populations vary during the run, and the initiator approximation is applied dynamically as walkers move back and forth between the unpopulated non-initiator, populated non-initiator, and initiator categories according to their instantaneous populations. The dynamic character of the initiator approximation prevents a straightforward algebraic analysis of the initiator bias; however, one can nonetheless construct a modified “initiator Hamiltonian” $\hat{\tilde{H}}$ that approximates its effects. We shall refer to $\hat{\tilde{H}} - \hat{H}$ as the “initiator perturbation”.

Let us assume knowledge of the exact eigenvector $\{C_i\}$, with determinants sorted so that $|C_i| \leq |C_j|$ if $i > j$, and let $C = \sum_i |C_i|$ be the $L1$ norm of the coefficients. Prior to performing an FCIQMC calculation with N walkers, one could then classify determinants as initiators if $|C_i| > N_{\text{init}}C/N$, as populated non-initiators if $C/N \leq |C_i| \leq N_{\text{init}}C/N$, and as unpopulated non-initiators if $|C_i| < N_{\text{pop}}C/N$, where $N_{\text{pop}} = 1/2$ would be a reasonable value. This classification defines three shells of determinants: shell 1 contains all n_1 initiator determinants, shell 2 contains all n_2 populated non-initiators, and shell 3 contains all n_3 unpopulated non-initiators, adding up to $n_1 + n_2 + n_3 = n_{\text{FCI}}$. The (symmetric) Hamiltonian matrix is consequently partitioned into nine blocks, and the Schrödinger equation is formulated as follows,

$$\begin{pmatrix} A & T & U \\ T^T & B & V \\ U^T & V^T & Z \end{pmatrix} \begin{pmatrix} \mathbf{c} \\ \mathbf{d} \\ \mathbf{g} \end{pmatrix} = E \begin{pmatrix} \mathbf{c} \\ \mathbf{d} \\ \mathbf{g} \end{pmatrix}, \quad (6.5)$$

where $\mathbf{c}$, $\mathbf{d}$, and $\mathbf{g}$ are vectors of the determinant coefficients in shells 1, 2, and 3, respectively, and the Hamiltonian matrix elements are

$$\begin{cases} A_{ji} = \langle 1_j | \hat{H} | 1_i \rangle \\ B_{ji} = \langle 2_j | \hat{H} | 2_i \rangle, \\ Z_{ji} = \langle 3_j | \hat{H} | 3_i \rangle \end{cases}, \quad \begin{cases} T_{ji} = \langle 1_j | \hat{H} | 2_i \rangle \\ U_{ji} = \langle 1_j | \hat{H} | 3_i \rangle, \\ V_{ji} = \langle 2_j | \hat{H} | 3_i \rangle \end{cases}, \quad (6.6)$$

where $|1_i\rangle$, $|2_i\rangle$, and $|3_i\rangle$ refer to the i th determinant in shells 1, 2, and 3, respectively. The dynamics enforced by the rules of the initiator approximation can be obtained by setting non-diagonal Hamiltonian matrix elements to zero between non-initiators and unpopulated

non-initiators, provided the determinant populations are set to their average values rounded to integers, $N_i = \lfloor NC_i/C \rfloor$. This amounts to defining the initiator Hamiltonian as

$$\hat{H} = \begin{pmatrix} A & T & U \\ T^T & B & 0 \\ U^T & 0 & \tilde{Z} \end{pmatrix}, \quad (6.7)$$

where $\tilde{Z}$ is a diagonal matrix given by $\tilde{Z}_{ij} = Z_{ij}\delta_{ij}$. Note that in real FCIQMC calculations, populations instantaneously differ from their average values, enabling “unpopulated” determinants to become momentarily occupied, and the initiator approximation allows these to spawn back onto occupied non-initiators. The initiator Hamiltonian $\tilde{H}$ does not capture this process, but as we will show later this does not affect our analysis.

The application of the unmodified FCIQMC master equation 6.2 to the initiator Hamiltonian can be expected to yield the initiator FCIQMC wave function and energy, and the effect of the initiator approximation can then be studied by comparing the solution of the Schrödinger equation with the initiator Hamiltonian,

$$\begin{pmatrix} A & T & U \\ T^T & B & 0 \\ U^T & 0 & \tilde{Z} \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{c}} \\ \tilde{\mathbf{d}} \\ \tilde{\mathbf{g}} \end{pmatrix} = \tilde{E} \begin{pmatrix} \tilde{\mathbf{c}} \\ \tilde{\mathbf{d}} \\ \tilde{\mathbf{g}} \end{pmatrix}, \quad (6.8)$$

with that of the Equation 6.5.

6.2.2 Scalar Measure of the Magnitude of the Initiator Perturbation

Characterizing the behavior of the initiator bias reduces to expressing $\tilde{E} - E$ as a function of the initiator perturbation, or more usefully, as a function of a scalar measure of its magnitude. Such a measure can be obtained by defining a suitable auxiliary system. Consider replacing all non-initiator determinants in the Hilbert space with a single site, $|3\rangle$, and requiring that the Schrödinger equation for the resulting system yield the same eigenvalue as for the full system,

$$\begin{pmatrix} A & \mathbf{t} \\ \mathbf{t}^T & b \end{pmatrix} \begin{pmatrix} \mathbf{c} \\ d \end{pmatrix} = E \begin{pmatrix} \mathbf{c} \\ d \end{pmatrix}, \quad (6.9)$$

where $\mathbf{t}$ is a vector of size n_1 , b and d are scalars, and A , $\mathbf{c}$, and E are defined to be the same as in Equation 6.5. Noting that d is not uniquely defined, we introduce an ansatz

that conveniently preserves the $L1$ norm of the coefficients,

$$d = \sum_{i=1}^{n_2} |d_i| + \sum_{i=1}^{n_3} |g_i|. \quad (6.10)$$

Equating terms in Equations 6.5 and 6.9 yields for the matrix elements

$$b = \frac{1}{d^2} \left(\sum_{i,j=1}^{n_2} B_{ij} d_i d_j + 2 \sum_{i=1}^{n_2} \sum_{j=1}^{n_3} V_{ji} d_i g_j + \sum_{i,j=1}^{n_3} Z_{ij} g_i g_j + E \left[d^2 - \sum_{i=1}^{n_2} d_i^2 - \sum_{i=1}^{n_3} g_i^2 \right] \right), \quad (6.11)$$

and

$$t_i = \frac{1}{d} \left(\sum_{j=1}^{n_2} T_{ji} d_j + \sum_{j=1}^{n_3} U_{ji} g_j \right). \quad (6.12)$$

Analogously, the auxiliary system for the initiator Hamiltonian satisfies the equation

$$\begin{pmatrix} A & \tilde{\mathbf{t}} \\ \tilde{\mathbf{t}}^T & \tilde{b} \end{pmatrix} \begin{pmatrix} \tilde{\mathbf{c}} \\ \tilde{d} \end{pmatrix} = \tilde{E} \begin{pmatrix} \tilde{\mathbf{c}} \\ \tilde{d} \end{pmatrix}, \quad (6.13)$$

with

$$\tilde{d} = \sum_{i=1}^{n_2} |\tilde{d}_i| + \sum_{i=1}^{n_3} |\tilde{g}_i|, \quad (6.14)$$

where

$$\tilde{b} = \frac{1}{\tilde{d}^2} \left(\sum_{i,j=1}^{n_2} B_{ij} \tilde{d}_i \tilde{d}_j + \sum_{i=1}^{n_3} Z_{ii} \tilde{g}_i^2 + \tilde{E} \left[\tilde{d}^2 - \sum_{i=1}^{n_2} \tilde{d}_i^2 - \sum_{i=1}^{n_3} \tilde{g}_i^2 \right] \right), \quad (6.15)$$

and

$$\tilde{t}_i = \frac{1}{\tilde{d}} \left(\sum_{j=1}^{n_2} T_{ji} \tilde{d}_j + \sum_{j=1}^{n_3} U_{ji} \tilde{g}_j \right). \quad (6.16)$$

Disregarding differences in eigenvector components, it is clear from Equations 6.11 and 6.15 that $\tilde{b} \approx b + \Delta$, where

$$\Delta = \frac{-1}{d^2} \left(2 \sum_{i=1}^{n_2} \sum_{j=1}^{n_3} V_{ji} d_i g_j + \sum_{i=1}^{n_3} \sum_{j \neq i}^{n_3} Z_{ij} g_i g_j \right). \quad (6.17)$$

As defined, Δ captures the direct effect of the initiator perturbation on the Hamiltonian of the auxiliary system, and can be used as a suitable scalar measure of the magnitude of the initiator perturbation in the full system.

Dependence of $\tilde{E}$ on Δ

Adding a small perturbation to a Hamiltonian causes first-order changes in its eigenvector and eigenvalue, so replacing b with $b + \Delta$ in the Hamiltonian of Equation 6.9 would shift the energy by an amount proportional to Δ . The resulting Hamiltonian does not coincide with that in Equation 6.13, because the definition of b and $\tilde{b}$ in the construction of the auxiliary systems involve the eigenvectors of the full systems, which differ; however, all definitions allow first-order expansions and the eigenvalue remains first-order in Δ , i.e., $(\tilde{E} - E) \propto \Delta$ at small Δ .

Dependence of Δ on N

It is convenient at this point to revert to referring to the unpartitioned Hamiltonian matrix elements, H_{ij} , and rearranging the definition of Δ of Equation 6.17 into

$$\Delta = \frac{-2}{d^2} \sum_{i=n_{\text{pop}}+1}^{n_{\text{FCI}}} \sum_{j=n_{\text{init}}+1}^{i-1} H_{ij} C_i C_j, \quad (6.18)$$

where $n_{\text{pop}} = n_1 + n_2$ and $n_{\text{init}} = n_1$ are the smallest integers that satisfy $|C_{n_{\text{pop}}+1}|N/C < N_{\text{pop}}$ and $|C_{n_{\text{init}}+1}|N/C < N_{\text{init}}$, respectively. For convenience we replace these definitions with the approximate $|C_{n_{\text{pop}}}| \approx N_{\text{pop}}C/N$ and $|C_{n_{\text{init}}}| \approx N_{\text{init}}C/N$. To make further progress, we need to make assumptions on the behavior of $|C_i|$ at large i and of $H_{ij}C_iC_j$ at large i and j . Empirically, determinant coefficients are found to follow a power law with determinant index i at large i , $|C_i| \propto i^{-\alpha}$ ($\alpha > 0$). Note that this makes n_{pop} and n_{init} proportional to $N^{1/\alpha}$.

Non-diagonal Hamiltonian matrix elements can be expected to connect only determinants which are in each other's "neighborhood", e.g., the Hamiltonian matrix element between a pair of determinants more than a double excitation apart in a molecular system is zero. Therefore, the sum $\sum_{j=n_{\text{init}}+1}^{i-1} H_{ij}C_iC_j$ can be expected to be dominated by terms where i and j are close in value, so we approximate it by

$$\sum_{j=n_{\text{init}}+1}^{i-1} H_{ij}C_iC_j \approx -h_iC_i^2. \quad (6.19)$$

Additionally we assume the neighboring-site transition strength h_i to be independent of i , $h_i \equiv h$, and we arrive at

$$\Delta \approx \frac{2h}{d^2} \sum_{i=n_{\text{pop}}+1}^{n_{\text{FCI}}} C_i^2 \propto \sum_{i=n_{\text{pop}}+1}^{n_{\text{FCI}}} i^{-2\alpha}. \quad (6.20)$$

The sum in Equation 6.20 can be put in terms of generalized Harmonic numbers $H_n^{(\gamma)} = \sum_{i=1}^n i^{-\gamma}$ as

$$\Delta \propto H_{n_{\text{FCI}}}^{(2\alpha)} - H_{n_{\text{pop}}}^{(2\alpha)}. \quad (6.21)$$

Generalized harmonic numbers asymptotically verify

$$H_n^{(\gamma)} \approx H_\infty^{(\gamma)} - \frac{n^{-(\gamma-1)}}{\gamma-1} + \mathcal{O}(n^{-\gamma}), \quad (6.22)$$

where $H_\infty^{(\gamma)}$ is a constant, and since n_{FCI} is extremely large we approximate $H_{n_{\text{FCI}}}^{(\gamma)} \approx H_\infty^{(\gamma)}$. Then we have

$$\Delta \propto n_{\text{pop}}^{-(2\alpha-1)}, \quad (6.23)$$

and taking into account that $n_{\text{pop}} \propto N^{1/\alpha}$ gives

$$\Delta \propto N^{-\beta}, \quad (6.24)$$

where $\beta = 2 - 1/\alpha$. The initiator bias therefore goes as $(\tilde{E} - E) \propto N^{-\beta}$ at large walker number N . The relationship between the exponent in the power law for determinant coefficients α and that governing the initiator bias β has certain interesting implications: (i) determinant coefficients must decay sufficiently fast, with $\alpha > 1/2$, for the initiator bias to follow a valid power law with $\beta > 0$; (ii) the initiator bias converges slowly, with $\beta < 1$, if $\alpha < 1$; (iii) the initiator bias converges with $\beta = 1$ if $\alpha = 1$; (iv) the initiator bias converges fast, with $\beta > 1$, if $\alpha > 1$; and (v) β tends to its limit value of 2 as $\alpha \rightarrow \infty$.

Current evidence points at $\beta \approx 1/3$ rather generally, implying that determinant coefficients should asymptotically follow a power law with exponent $\alpha = 3/5 = 0.6$. In Figure 6.1, we show the initiator bias as a function of $N^{-1/3}$ for basis sets of the cc-pVnZ family⁴⁰ ranging from double- to quintuple-zeta.

6.3 Results and Discussion

We apply our extrapolation method to the ground state of C_2 , CN , N_2 , and O_2 molecules in their equilibrium geometry¹ (Table 6.1) by performing all-electron FCIQMC calculations using Hartree-Fock orbitals expanded in the cc-pVnZ Gaussian basis sets. The orbitals are generated using the PYSCF package¹⁵⁴, and the FCIQMC calculations are performed with the NECI package^{12,80}.

In order to obtain benchmark energies to compare FCIQMC results against, we perform semistochastic heat-bath CI calculations using the DICE code^{22,55}. For the SHCI method, it

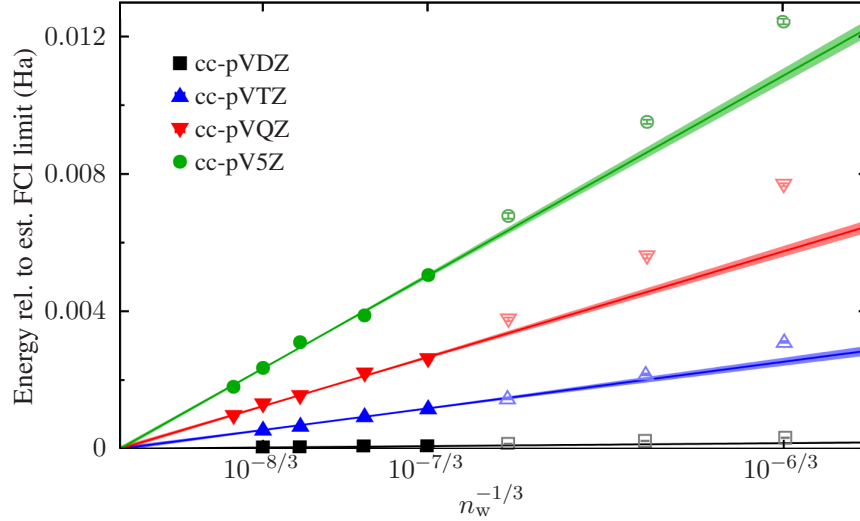


Figure 6.1: Estimated initiator error in the FCIQMC energies as a function of $n_w^{-1/3}$ for the CN molecule. The hollow points are not used in the extrapolations. The translucent areas represent the uncertainty in the values of the fit at each point.

	State	r_{eq} (Å)
C ₂	$1\Sigma_g^+$	1.2425
CN	$2\Sigma^+$	1.1718
N ₂	$1\Sigma_g^+$	1.0977
O ₂	$3\Sigma_g^-$	1.2075

Table 6.1: Electronic ground states and equilibrium bond lengths used for the first-row dimers considered in this work, following Ref. 1.

has been shown that one can obtain accurate estimates of total energies in the FCI limit by linear extrapolation of SHCI energies to the limit in which the second-order perturbation-theory (PT) correction E_2 is zero^{155,156}. The calculations are done with variational spaces of sizes such that $|E_2|$ is not much larger than about 10 mHa^{155,156}; we only exceed this threshold for the cc-pV5Z calculations due to memory constraints.

The FCIQMC calculations are performed using between $n_w = 10^6$ and 2×10^8 walkers. The results are reported in Table 6.2. The extrapolated energies are in excellent alignment, with FCI limits obtained from FCIQMC differing by 0.4 mHa or less from their SHCI counterparts throughout. Figure 6.2 illustrates this comparison for the CN molecule in

	Basis set	C ₂	CN	N ₂	O ₂
SHCI	cc-pVDZ	-75.7320	-92.4970	-109.2809	-149.9915
	cc-pVTZ	-75.8094	-92.5952	-109.4015	-150.1556
	cc-pVQZ	-75.8577	-92.6516	-109.4655	-150.2366
	cc-pV5Z	-75.8751	-92.6718	-109.4884	-150.2659
FCIQMC	cc-pVDZ	-75.7320	-92.4970	-109.2809	-149.9915
	cc-pVTZ	-75.8094	-92.5954	-109.4014	-150.1554
	cc-pVQZ	-75.8578	-92.6517	-109.4653	-150.2362
	cc-pV5Z	-75.8752	-92.6717	-109.4881	-150.2655

Table 6.2: Total energies of relevant dimers in the FCI limit in Ha, obtained by extrapolation of SHCI energies to the zero second-order PT correction limit and by extrapolation of FCIQMC energies to the infinite-walker limit.

the cc-pVnZ ($n=D, T, Q$, and 5) basis sets. Note that in these extrapolations, we do not include results at $n_w < 10^7$ walkers, which we deem to lie outside of the asymptotic regime.

To further test our method, we consider the ground state of the chromium atom (7S) as well as the ground state of the highly challenging chromium dimer (${}^1\Sigma_g^+$) in its equilibrium geometry ($r_{eq} = 1.68 \text{ \AA}$ ¹⁵⁷). We perform frozen-core FCIQMC calculations with the unfrozen $3sp$ states, using the Hartree-Fock orbitals expanded in the cc-pVnZ-DK ($n=D, T$, and Q) basis sets ¹⁵⁸ generated with the MOLPRO package ^{44,45}. Relativistic effects are included using the X2C Hamiltonian ¹⁵⁹ for the Cr₂ dimer, and using the Douglas-Kroll-Hess ^{160,161} and X2C Hamiltonians for the Cr atom.

In Figure 6.3, we plot the FCIQMC energy (with the X2C Hamiltonian) relative to its corresponding FCI limit, i.e., the initiator error, as a function of $n_w^{-1/3}$ for the Cr atom and dimer for each of the basis sets considered. The extrapolations are done using the walker population ranges from $n_w = 6 \times 10^7$ to 7.2×10^8 and from $n_w = 6 \times 10^7$ to 4.3×10^9 walkers for the Cr atom and dimer, respectively.

Our extrapolated total and relative (atomization) energies are tabulated in Table 6.3, along with highly-accurate results from the literature. Our total energies for the Cr atom with the Douglas-Kroll-Hess Hamiltonian are in excellent agreement with the large-population FCIQMC results obtained using natural orbitals. The total energies go down with increasing basis-set size, but surprisingly the atomization energy obtained with the DZ basis is larger than those obtained using the TZ and QZ basis sets. This is likely due to a larger basis-set superposition error (BSSE) in the DZ basis compared to the TZ and QZ

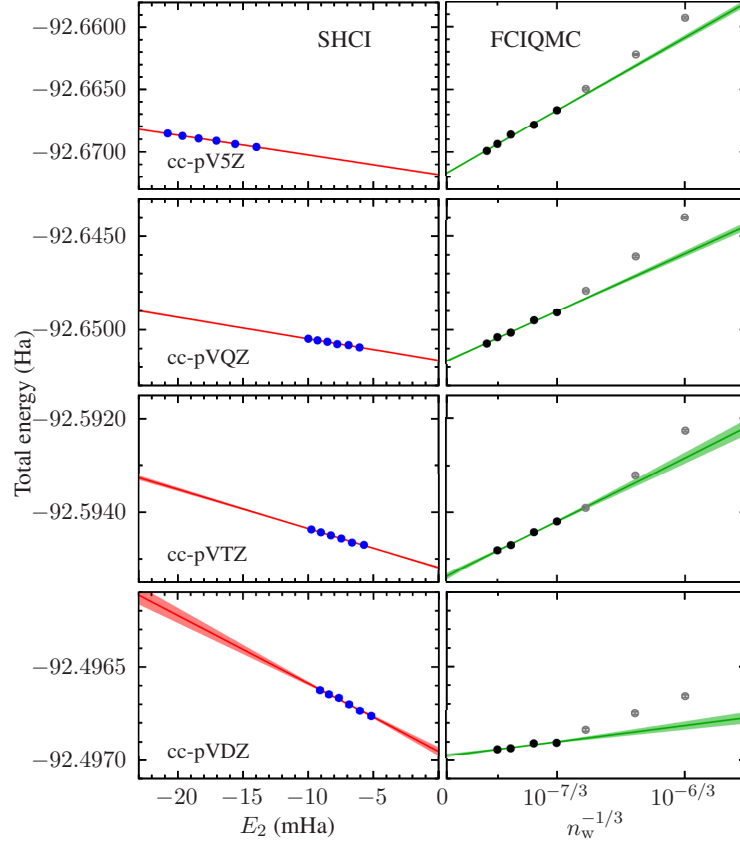


Figure 6.2: Extrapolation to the FCI limit of SHCI energies with respect to the second-order PT correction E_2 (left), and of FCIQMC energies with respect to $n_w^{-1/3}$ (right) for the CN molecule. The grayed-out points are not used in the FCIQMC extrapolations. The width of the line reflects the statistical uncertainty within the fit.

bases. This bias can be eliminated using the counterpoise (CP) method. Also, the level of convergence of the extrapolated energy for the Cr_2 dimer up to $n_w = 4.3 \times 10^9$ walkers is unsatisfactory. The four-billion-walker run is a very expensive calculation for the Cr_2 dimer; however, to achieve a higher level of convergence, one needs to run larger-walker-number calculations, particularly in the QZ basis.

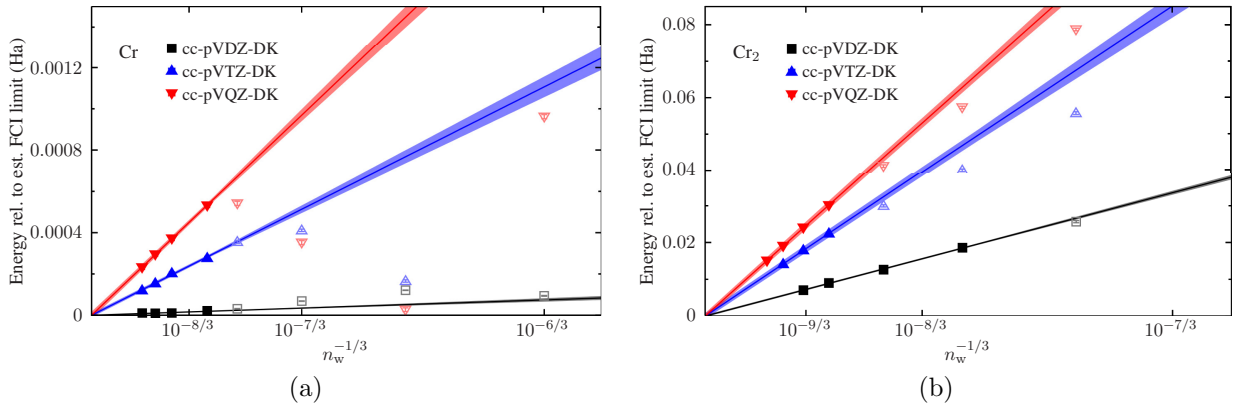


Figure 6.3: Estimated initiator error in the FCIQMC energies as a function of $n_w^{-1/3}$ for the chromium atom and dimer with the X2C relativistic Hamiltonian. The empty symbols denote points not used in the extrapolations. The line width represents the statistical uncertainty on the fit values at each point.

	Basis set	Cr	Cr ₂	Atomization
SHCI ^a	cc-pVDZ-DK			0.05724
	cc-pVTZ-DK			0.04812
	cc-pVQZ-DK			0.05530
FCIQMC ^b	cc-pVDZ-DK	−1049.925266(1)	−2099.90240(1)	0.05187(1)
	cc-pVTZ-DK	−1050.081559(4)	−2100.20299(6)	0.03987(6)
	cc-pVQZ-DK	−1050.144276(11)	−2100.34178(6)	0.05323(6)
This work ^b	cc-pVDZ-DK	−1049.925313(2)		
	cc-pVTZ-DK	−1050.081926(9)		
	cc-pVQZ-DK	−1050.144769(16)		
This work ^a	cc-pVDZ-DK	−1049.932568(2)	−2099.91956(27)	0.05442(27)
	cc-pVTZ-DK	−1050.089181(10)	−2100.21844(80)	0.04008(80)
	cc-pVQZ-DK	−1050.152122(17)	−2100.34766(78)	0.04342(78)

^aWith the X2C relativistic Hamiltonian.

^bWith the Douglas-Kroll-Hess relativistic Hamiltonian.

Table 6.3: Total energies of the chromium atom and dimer, and atomization energies of the chromium dimer, in Ha. Results labeled “SHCI” are from Ref. 2, those labeled “FCIQMC” are large-population FCIQMC results obtained using natural orbitals from Ref. 3, and those labeled “This work” are FCIQMC energies obtained with Hartree-Fock orbitals extrapolated to the infinite-walker limit.

Part IV

Summary and Outlook

In this thesis, our main focus was on the combination of orbital and real space quantum Monte Carlo methods. In the first chapter, the electronic structure problem was formulated, where we introduced some basic concepts and definitions in quantum chemistry. Starting from the many-body Schrödinger equation, we then discussed a general statement of the Born-Oppenheimer approximation, molecular orbitals, the antisymmetry principle, and second quantization. Some fundamental aspects of periodic solids including the Bravais lattice and twist averaging along with an introduction to the three-dimensional uniform electron gas were also discussed. In the next two chapters, we made a review on the two main groups of *ab initio* methods, namely quantum chemistry and quantum Monte Carlo. Hartree-Fock theory and the configuration interaction method were introduced as the most important quantum chemistry approaches to the calculation of the electronic structure. Following this, an in-depth review was provided for various QMC methods, including VMC, DMC, and FCIQMC. Moreover, the transcorrelated FCIQMC method was explored.

After introducing the theory, we first presented an empirical extrapolation method for continuum QMC energies as a function of the sum of the squared coefficients in the initial selected CI wave function from which the trial wave function is constructed. This approach is made possible by the smoothness of the energies as a function of the CI expansion size, and we presented a simple statistical procedure to handle the quasi-random non-smoothness in the data, which we showed to work very well in practice. We found that HF orbitals expanded in standard basis sets provide the type of gradual convergence required for the XSPOT method to work well. The results from the tests we conducted are compatible with the XSPOT method being capable of obtaining exact total energies, with the caveat that trial wave function complexity must increase with system size in order to control the uncertainty on the results. The application of this method to a wider variety of systems and attempting to scale the method with system size with controlled uncertainties will be interesting topics for future studies.

Afterwards, we extrapolated the fixed-node DMC correlation energy and fixed-node error of the high-density spin-unpolarized uniform electron gas to the thermodynamic limit, and used this in combination with TC-FCIQMC results applying the method proposed in Ref. 21 to estimate the thermodynamic limit of the exact correlation energy at $r_s = 0.5$, which is found to lie within statistical uncertainty of the QMC correlation energy recently reported at this density. However, the fixed-node errors were found to behave non-monotonically as a function of the system size N , leading to a relatively large statistical uncertainty for the extrapolated thermodynamic limit. Given that, future efforts can be focused on obtaining more accurate TC-FCIQMC correlation energies using our newly proposed walker number extrapolation technique.

In the last chapter, we developed a new model to treat the energy bias coming from the initiator approximation in the FCIQMC method. We found that extrapolating the FCIQMC energy linearly in $n_w^{-1/3}$ to $n_w^{-1/3} \rightarrow 0$ eliminates the initiator error and converges the energy to the full configuration interaction limit. This type of extrapolation can be formally justified by defining an “initiator perturbation” and modelling the initiator bias as a scalar measure of its magnitude. The application of this extrapolation scheme to a series of dimers yielded results in excellent agreement with the benchmark FCI limits from the semistochastic heat-bath CI method. In the future, we will investigate extending the formal justification to provide the value of the exponent of the power law ($-1/3$) instead of its general form.

Appendix A

The Slater-Condon Rules

The electronic Hamiltonian 1.6 is the sum of one- and two-body operators,

$$\hat{H} = \hat{\mathcal{O}}_1 + \hat{\mathcal{O}}_2 = \sum_{i=1}^N \hat{h}(i) + \sum_{i=1}^N \sum_{j>i}^N r_{ij}^{-1}. \quad (25)$$

Given two determinants $|D_i\rangle$ and $|D_j\rangle$, the Hamiltonian matrix element between these two determinants is as follows,

$$\langle D_i | \hat{H} | D_j \rangle = \langle D_i | \hat{\mathcal{O}}_1 | D_j \rangle + \langle D_i | \hat{\mathcal{O}}_2 | D_j \rangle, \quad (26)$$

where the first and the second terms yield one- and two-body integrals, respectively. The two determinants are said to be coupled if and only if $\langle D_i | \hat{H} | D_j \rangle \neq 0$. Because of the one- and two-body nature of the electronic Hamiltonian, every two determinants which differ by more than two (orthonormal) spin orbitals are then uncoupled. Thus matrix elements are just computed for determinants which are single or double excitations of each other. Such matrix elements are computed using the Slater-Condon rules^{36–38}. When the two determinants are the same, the diagonal matrix elements, using the chemists' notation for one- and two-body integrals over spin orbitals, are given by

$$\langle D_i | \hat{H} | D_i \rangle = \sum_a^N [a | \hat{h} | a] + \sum_a^N \sum_{b>a}^N [aa | bb] - [ab | ba]. \quad (27)$$

In the case of single excitations, let $|D_j\rangle = \hat{a}_r^\dagger \hat{a}_p |D_i\rangle$, then

$$\langle D_i | \hat{H} | D_j \rangle = [p | \hat{h} | r] + \sum_a^N [pr | aa] - [pa | ar]. \quad (28)$$

Also for double excitations, let $|D_j\rangle = \hat{a}_r^t \hat{a}_s^t \hat{a}_q \hat{a}_p |D_i\rangle$, then the matrix elements would be

$$\langle D_i | \hat{H} | D_j \rangle = [pr|qs] - [ps|qr]. \quad (29)$$

The transcorrelated (TC) method (see [3.5.1](#)) modifies the electronic Hamiltonian by requiring three-body terms (refer to Equations [3.60](#) and [3.61](#)), which are handled separately from the above rules [98,102](#).

References

- [1] D. Feller, K. A. Peterson, and D. A. Dixon, A survey of factors contributing to accurate theoretical predictions of atomization energies and molecular structures, [J. Chem. Phys. **129**, 204105 \(2008\)](#).
- [2] J. Li, Y. Yao, A. A. Holmes, M. Otten, Q. Sun, S. Sharma, and C. J. Umrigar, Accurate many-body electronic structure near the basis set limit: application to the chromium dimer, [Phys. Rev. Research **2**, 012015\(R\) \(2020\)](#).
- [3] A. Y. Lozovoi, S. Sharma, and A. Alavi, Bonding and the role of semicore in chromium dimer: a combined study by full configuration interaction quantum Monte Carlo and linearized coupled cluster methods, Unpublished (2017).
- [4] D. M. Ceperley, Fermion nodes, [J. Stat. Phys. **63**, 1237 \(1991\)](#).
- [5] L. Bytautas and K. Ruedenberg, Correlation energy extrapolation by intrinsic scaling. IV. Accurate binding energies of the homonuclear diatomic molecules carbon, nitrogen, oxygen, and fluorine, [J. Chem. Phys. **122**, 154110 \(2005\)](#).
- [6] P. Hohenberg and W. Kohn, Inhomogeneous electron gas, [Phys. Rev. **136**, B864 \(1964\)](#).
- [7] W. Kohn and L. J. Sham, Self-consistent equations including exchange and correlation effects, [Phys. Rev. **140**, A1133 \(1965\)](#).
- [8] A. Szabo and N. S. Ostlund, *Modern Quantum Chemistry* (McGraw-Hill, New York, 1996).
- [9] T. Helgaker, P. Jørgensen, and J. Olsen, *Molecular Electronic-Structure Theory* (Wiley, Chichester, 2000).
- [10] W. L. McMillan, Ground state of liquid He^4 , [Phys. Rev. **138**, A442 \(1965\)](#).

- [11] D. M. Ceperley and B. J. Alder, Ground state of the electron gas by a stochastic method, [Phys. Rev. Lett. **45**, 566 \(1980\)](#).
- [12] G. H. Booth, A. J. W. Thom, and A. Alavi, Fermion Monte Carlo without fixed nodes: a game of life, death, and annihilation in Slater determinant space, [J. Chem. Phys. **131**, 054106 \(2009\)](#).
- [13] D. Cleland, G. H. Booth, and A. Alavi, Communications: survival of the fittest: accelerating convergence in full configuration-interaction quantum Monte Carlo, [J. Chem. Phys. **132**, 041103 \(2010\)](#).
- [14] R. J. Needs, M. D. Towler, N. D. Drummond, P. López Ríos, and J. R. Trail, Variational and diffusion quantum Monte Carlo calculations with the CASINO code, [J. Chem. Phys. **152**, 154106 \(2020\)](#).
- [15] N. D. Drummond, M. D. Towler, and R. J. Needs, Jastrow correlation factor for atoms, molecules, and solids, [Phys. Rev. B **70**, 235119 \(2004\)](#).
- [16] P. López Ríos, P. Seth, N. D. Drummond, and R. J. Needs, Framework for constructing generic Jastrow correlation factors, [Phys. Rev. E **86**, 036703 \(2012\)](#).
- [17] P. López Ríos, A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs, Inhomogeneous backflow transformations in quantum Monte Carlo calculations, [Phys. Rev. E **74**, 066701 \(2006\)](#).
- [18] J. S. Spencer, N. S. Blunt, and W. M. C. Foulkes, The sign problem and population dynamics in the full configuration interaction quantum Monte Carlo method, [J. Chem. Phys. **136**, 054110 \(2012\)](#).
- [19] M. H. Kolodrubetz, J. S. Spencer, B. K. Clark, and W. M. C. Foulkes, The effect of quantization on the full configuration interaction quantum Monte Carlo sign problem, [J. Chem. Phys. **138**, 024110 \(2013\)](#).
- [20] P.-F. Loos and P. M. W. Gill, The uniform electron gas, [Wiley Interdiscip. Rev.: Comput. Mol. Sci. **6**, 410 \(2016\)](#).
- [21] M. Ruggeri, P. López Ríos, and A. Alavi, Correlation energies of the high-density spin-polarized electron gas to meV accuracy, [Phys. Rev. B **98**, 161105\(R\) \(2018\)](#).
- [22] S. Sharma, A. A. Holmes, G. Jeanmairet, A. Alavi, and C. J. Umrigar, Semistochastic heat-bath configuration interaction method: selected configuration interaction with semistochastic perturbation theory, [J. Chem. Theory Comput. **13**, 1595 \(2017\)](#).

- [23] M. Born and R. Oppenheimer, On the quantum theory of molecules, [Ann. Phys. **389**, 457 \(1927\)](#).
- [24] J. C. Tully, Nonadiabatic Processes in Molecular Collisions, in *Dynamics of Molecular Collisions*, edited by W. H. Miller (Plenum Press, New York, 1976).
- [25] J. Paldus, Group theoretical approach to the configuration interaction and perturbation theory calculations for atomic and molecular systems, [J. Chem. Phys. **61**, 5321 \(1974\)](#).
- [26] I. M. Gel'fand and M. L. Tsetlin, Finite-dimensional representations of the group of unimodular matrices, Doklady Akad. Nauk SSSR (N.S.) **71**, 825 (1950).
- [27] I. M. Gel'fand and M. L. Tsetlin, Finite-dimensional representations of the group of orthogonal matrices, Doklady Akad. Nauk SSSR (N.S.) **71**, 1017 (1950) [Am. Math. Soc. Transl. **64**, 116 (1967)].
- [28] I. M. Gel'fand, The center of an infinitesimal group ring, Mat. Sb. (N.S.) **26**(68), 103 (1950).
- [29] I. Shavitt, The Graphical Unitary Group Approach and Its Application to Direct Configuration Interaction Calculations, in *The Unitary Group for the Evaluation of Electronic Energy Matrix Elements*, edited by J. Hinze (Springer-Verlag Berlin, Heidelberg, 1981).
- [30] H. Weyl, *The Classical Groups: Their Invariants and Representations* (Princeton University Press, Princeton, 1946).
- [31] G. F. Giuliani and G. Vignale, *Quantum Theory of the Electron Liquid* (Cambridge University Press, New York, 2005).
- [32] E. Wigner, On the interaction of electrons in metals, [Phys. Rev. **46**, 1002 \(1934\)](#).
- [33] C. Lin, F. H. Zong, and D. M. Ceperley, Twist-averaged boundary conditions in continuum quantum Monte Carlo algorithms, [Phys. Rev. E **64**, 016702 \(2001\)](#).
- [34] S. Chiesa, D. M. Ceperley, R. M. Martin, and M. Holzmann, Finite-size error in many-body simulations with long-range interactions, [Phys. Rev. Lett. **97**, 076404 \(2006\)](#).
- [35] N. D. Drummond, R. J. Needs, A. Sorouri, and W. M. C. Foulkes, Finite-size errors in continuum quantum Monte Carlo calculations, [Phys. Rev. B **78**, 125106 \(2008\)](#).

- [36] J. C. Slater, The theory of complex spectra, [Phys. Rev. **34**, 1293 \(1929\)](#).
- [37] E. U. Condon, The theory of complex spectra, [Phys. Rev. **36**, 1121 \(1930\)](#).
- [38] J. C. Slater, Molecular energy levels and valence bonds, [Phys. Rev. **38**, 1109 \(1931\)](#).
- [39] P. E. Cade and A. C. Wahl, Hartree-Fock-Roothaan wavefunctions for diatomic molecules: II. First-row homonuclear systems A_2 , $A_2^\pm$, and A_2^* , [At. Data Nucl. Data Tables **13**, 339 \(1974\)](#).
- [40] T. H. Dunning Jr., Gaussian basis sets for use in correlated molecular calculations. I. The atoms boron through neon and hydrogen, [J. Chem. Phys. **90**, 1007 \(1989\)](#).
- [41] D. E. Woon and T. H. Dunning Jr., Gaussian basis sets for use in correlated molecular calculations. V. Core-valence basis sets for boron through neon, [J. Chem. Phys. **103**, 4572 \(1995\)](#).
- [42] A. K. Wilson, T. van Mourik, and T. H. Dunning Jr, Gaussian basis sets for use in correlated molecular calculations. VI. Sextuple zeta correlation consistent basis sets for boron through neon, [J. Mol. Struct.: THEOCHEM **388**, 339 \(1996\)](#).
- [43] A. J. C. Varandas, CBS extrapolation of Hartree-Fock energy: Pople and Dunning basis sets hand-to-hand on the endeavour, [Phys. Chem. Chem. Phys. **21**, 8022 \(2019\)](#).
- [44] H.-J. Werner, P. J. Knowles, G. Knizia, F. R. Manby, and M. Schütz, Molpro: a general-purpose quantum chemistry program package, [Wiley Interdiscip. Rev.: Comput. Mol. Sci. **2**, 242 \(2012\)](#).
- [45] H.-J. Werner, P. J. Knowles, F. R. Manby, J. A. Black, K. Doll, A. Heßelmann, D. Kats, A. Köhn, T. Korona, D. A. Kreplin, Q. Ma, T. F. Miller III, A. Mitrushchenkov, K. A. Peterson, I. Polyak, G. Rauhut, and M. Sibaev, The Molpro quantum chemistry package, [J. Chem. Phys. **152**, 144107 \(2020\)](#).
- [46] J. Ivanic and K. Ruedenberg, Identification of deadwood in configuration spaces through general direct configuration interaction, [Theor. Chem. Acc. **106**, 339 \(2001\)](#).
- [47] B. Huron, J. P. Malrieu, and P. Rancurel, Iterative perturbation calculations of ground and excited state energies from multiconfigurational zeroth-order wavefunctions, [J. Chem. Phys. **58**, 5745 \(1973\)](#).
- [48] S. Evangelisti, J. P. Daudey, and J. P. Malrieu, Convergence of an improved CIPSI algorithm, [Chem. Phys. **75**, 91 \(1983\)](#).

- [49] R. J. Harrison, Approximating full configuration interaction with selected configuration interaction and perturbation theory, [J. Chem. Phys. **94**, 5021 \(1991\)](#).
- [50] M. L. Abrams and C. D. Sherrill, Important configurations in configuration interaction and coupled-cluster wave functions, [Chem. Phys. Lett. **412**, 121 \(2005\)](#).
- [51] P. J. Knowles, Compressive sampling in configuration interaction wavefunctions, [Mol. Phys. **113**, 1655 \(2015\)](#).
- [52] M. Caffarel, T. Applencourt, E. Giner, and A. Scemama, Using CIPSI Nodes in Diffusion Monte Carlo, in *Recent Progress in Quantum Monte Carlo*, edited by S. Tanaka, P.-N. Roy, and L. Mitas (ACS Publications, Washington, DC, 2016), Chap. 2.
- [53] J. B. Schriber and F. A. Evangelista, Communication: an adaptive configuration interaction approach for strongly correlated electrons with tunable accuracy, [J. Chem. Phys. **144**, 161106 \(2016\)](#).
- [54] N. M. Tubman, J. Lee, T. Y. Takeshita, M. Head-Gordon, and K. B. Whaley, A deterministic alternative to the full configuration interaction quantum Monte Carlo method, [J. Chem. Phys. **145**, 044112 \(2016\)](#).
- [55] A. A. Holmes, N. M. Tubman, and C. J. Umrigar, Heat-bath configuration interaction: an efficient selected configuration interaction algorithm inspired by heat-bath sampling, [J. Chem. Theory Comput. **12**, 3674 \(2016\)](#).
- [56] Y. Garniron, A. Scemama, E. Giner, M. Caffarel, and P.-F. Loos, Selected configuration interaction dressed by perturbation, [J. Chem. Phys. **149**, 064103 \(2018\)](#).
- [57] P. S. Epstein, The Stark effect from the point of view of Schrödinger's quantum theory, [Phys. Rev. **28**, 695 \(1926\)](#).
- [58] R. K. Nesbet, Configuration interaction in orbital theories, [Proc. R. Soc. London, Ser. A **230**, 312 \(1955\)](#).
- [59] N. Metropolis and S. Ulam, The Monte Carlo method, [J. Am. Stat. Assoc. **44**, 335 \(1949\)](#).
- [60] N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, Equation of state calculations by fast computing machines, [J. Chem. Phys. **21**, 1087 \(1953\)](#).

- [61] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Quantum Monte Carlo simulations of solids, *Rev. Mod. Phys.* **73**, 33 (2001).
- [62] W. M. C. Foulkes, R. Q. Hood, and R. J. Needs, Symmetry constraints and variational principles in diffusion quantum Monte Carlo calculations of excited-state energies, *Phys. Rev. B* **60**, 4558 (1999).
- [63] M. H. Kalos, Monte Carlo calculations of the ground state of three- and four-body nuclei, *Phys. Rev.* **128**, 1791 (1962).
- [64] M. H. Kalos, D. Levesque, and L. Verlet, Helium at zero temperature with hard-sphere and other forces, *Phys. Rev. A* **9**, 2178 (1974).
- [65] J. B. Anderson, A random-walk simulation of the Schrödinger equation: H_3^+ , *J. Chem. Phys.* **63**, 1499 (1975).
- [66] J. B. Anderson, Quantum chemistry by random walk. $\text{H } ^2P$, $\text{H}_3^+ D_{3h} ^1A'_1$, $\text{H}_2 ^3\Sigma_u^+$, $\text{H}_4 ^1\Sigma_g^+$, $\text{Be } ^1S$, *J. Chem. Phys.* **65**, 4121 (1976).
- [67] P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester, Fixed-node quantum Monte Carlo for molecules, *J. Chem. Phys.* **77**, 5593 (1982).
- [68] J. W. Moskowitz, K. E. Schmidt, M. A. Lee, and M. H. Kalos, A new look at correlation energy in atomic and molecular systems. II. The application of the Green's function Monte Carlo method to LiH , *J. Chem. Phys.* **77**, 349 (1982).
- [69] R. C. Grimm and R. G. Storer, Monte-Carlo solution of Schrödinger's equation, *J. Comput. Phys.* **7**, 134 (1971).
- [70] R. M. Lee, G. J. Conduit, N. Nemec, P. López Ríos, and N. D. Drummond, Strategies for improving the efficiency of quantum Monte Carlo calculations, *Phys. Rev. E* **83**, 066706 (2011).
- [71] P. López Ríos (2023) POLYFIT [Source code]. <https://github.com/plopezrios/poly-fit>.
- [72] R. J. Needs, M. D. Towler, N. D. Drummond, and P. López Ríos, Continuum variational and diffusion quantum Monte Carlo calculations, *J. Phys.: Condens. Matter* **22**, 023201 (2010).
- [73] C. J. Umrigar, M. P. Nightingale, and K. J. Runge, A diffusion Monte Carlo algorithm with very small time-step errors, *J. Chem. Phys.* **99**, 2865 (1993).

- [74] R. Jastrow, Many-body problem with strong forces, *Phys. Rev.* **98**, 1479 (1955).
- [75] T. Kato, On the eigenfunctions of many-particle systems in quantum mechanics, *Commun. Pure Appl. Math.* **10**, 151 (1957).
- [76] R. T. Pack and W. B. Brown, Cusp conditions for molecular wavefunctions, *J. Chem. Phys.* **45**, 556 (1966).
- [77] C.-J. Huang, C. Filippi, and C. J. Umrigar, Spin contamination in quantum Monte Carlo wave functions, *J. Chem. Phys.* **108**, 8838 (1998).
- [78] R. P. Feynman, Atomic theory of the two-fluid model of liquid helium, *Phys. Rev.* **94**, 262 (1954).
- [79] R. P. Feynman and M. Cohen, Energy spectrum of the excitations in liquid helium, *Phys. Rev.* **102**, 1189 (1956).
- [80] K. Guthrie, R. J. Anderson, N. S. Blunt, N. A. Bogdanov, D. Cleland, N. Dattani, W. Dobbrautz, K. Ghanem, P. Jeszenszki, N. Liebermann, G. Li Manni, A. Y. Lozovoi, H. Luo, D. Ma, F. Merz, C. Overy, M. Rampp, P. K. Samanta, L. R. Schwarz, J. J. Shepherd, S. D. Smart, E. Vitale, O. Weser, G. H. Booth, and A. Alavi, NECI: *N*-electron configuration interaction with an emphasis on state-of-the-art stochastic methods, *J. Chem. Phys.* **153**, 034107 (2020).
- [81] S. F. Boys, Some bilinear convergence characteristics of the solutions of dissymmetric secular equations, *Proc. R. Soc. London, Ser. A* **309**, 195 (1969).
- [82] S. F. Boys and N. C. Handy, A condition to remove the indeterminacy in interelectronic correlation functions, *Proc. R. Soc. London, Ser. A* **309**, 209 (1969).
- [83] S. F. Boys and N. C. Handy, The determination of energies and wavefunctions with full electronic correlation, *Proc. R. Soc. London, Ser. A* **310**, 43 (1969).
- [84] S. F. Boys and N. C. Handy, A calculation for the energies and wavefunctions for states of neon with full electronic correlation accuracy, *Proc. R. Soc. London, Ser. A* **310**, 63 (1969).
- [85] S. F. Boys and N. C. Handy, A first solution, for LiH, of a molecular transcorrelated wave equation by means of restricted numerical integration, *Proc. R. Soc. London, Ser. A* **311**, 309 (1969).

- [86] N. C. Handy, Energies and expectation values for Be by the transcorrelated method, *J. Chem. Phys.* **51**, 3205 (1969).
- [87] N. C. Handy, On the minimization of the variance of the transcorrelated Hamiltonian, *Mol. Phys.* **21**, 817 (1971).
- [88] N. C. Handy, The transcorrelated method for accurate correlation energies using gaussian-type functions: examples on He, H₂, LiH and H₂O, *Mol. Phys.* **23**, 1 (1972).
- [89] N. C. Handy, Towards an understanding of the form of correlated wavefunctions for atoms, *J. Chem. Phys.* **58**, 279 (1973).
- [90] S. Ten-no, A feasible transcorrelated method for treating electronic cusps using a frozen Gaussian geminal, *Chem. Phys. Lett.* **330**, 169 (2000).
- [91] S. Ten-no, Three-electron integral evaluation in the transcorrelated method using a frozen Gaussian geminal, *Chem. Phys. Lett.* **330**, 175 (2000).
- [92] O. Hino, Y. Tanimura, and S. Ten-no, Application of the transcorrelated Hamiltonian to the linearized coupled cluster singles and doubles model, *Chem. Phys. Lett.* **353**, 317 (2002).
- [93] N. Umezawa and S. Tsuneyuki, Transcorrelated method for electronic systems coupled with variational Monte Carlo calculation, *J. Chem. Phys.* **119**, 10015 (2003).
- [94] N. Umezawa, S. Tsuneyuki, T. Ohno, K. Shiraishi, and T. Chikyow, A practical treatment for the three-body interactions in the transcorrelated variational Monte Carlo method: application to atoms from lithium to neon, *J. Chem. Phys.* **122**, 224101 (2005).
- [95] H. Luo and A. Alavi, Combining the transcorrelated method with full configuration interaction quantum Monte Carlo: application to the homogeneous electron gas, *J. Chem. Theory Comput.* **14**, 1403 (2018).
- [96] A. J. Cohen, H. Luo, K. Guthrie, W. Dobbrautz, D. P. Tew, and A. Alavi, Similarity transformation of the electronic Schrödinger equation via Jastrow factorization, *J. Chem. Phys.* **151**, 061101 (2019).
- [97] W. Dobbrautz, H. Luo, and A. Alavi, Compact numerical solutions to the two-dimensional repulsive Hubbard model obtained via nonunitary similarity transformations, *Phys. Rev. B* **99**, 075119 (2019).

- [98] J. P. Haupt, S. M. Hosseini, P. López Ríos, W. Dobrautz, A. Cohen, and A. Alavi, Optimizing Jastrow factors for the transcorrelated method, *J. Chem. Phys.* **158**, 224105 (2023).
- [99] T. Schraivogel, A. J. Cohen, A. Alavi, and D. Kats, Transcorrelated coupled cluster methods, *J. Chem. Phys.* **155**, 191101 (2021).
- [100] K. Liao, T. Schraivogel, H. Luo, D. Kats, and A. Alavi, Towards efficient and accurate *ab initio* solutions to periodic systems via transcorrelation and coupled cluster theory, *Phys. Rev. Res.* **3**, 033072 (2021).
- [101] T. Schraivogel, E. M. Christlmaier, P. López Ríos, A. Alavi, and D. Kats, Transcorrelated coupled cluster methods. II. Molecular systems, *J. Chem. Phys.* **158**, 214106 (2023).
- [102] E. M. Christlmaier, T. Schraivogel, P. López Ríos, A. Alavi, and D. Kats, xTC: an efficient treatment of three-body interactions in transcorrelated methods, *J. Chem. Phys.* **159**, 014113 (2023).
- [103] A. Baiardi and M. Reiher, Transcorrelated density matrix renormalization group, *J. Chem. Phys.* **153**, 164115 (2020).
- [104] K. Liao, H. Zhai, E. M. Christlmaier, T. Schraivogel, P. López Ríos, D. Kats, and A. Alavi, Density matrix renormalization group for transcorrelated Hamiltonians: ground and excited states in molecules, *J. Chem. Theory Comput.* **19**, 1734 (2023).
- [105] A. Ammar, A. Scemama, and E. Giner, Extension of selected configuration interaction for transcorrelated methods, *J. Chem. Phys.* **157**, 134107 (2022).
- [106] A. Ammar, A. Scemama, and E. Giner, Transcorrelated selected configuration interaction in a bi-orthonormal basis and with a cheap three-body correlation factor, *J. Chem. Phys.* **159**, 114121 (2023).
- [107] I. O. Sokolov, W. Dobrautz, H. Luo, A. Alavi, and I. Tavernelli, Orders of magnitude increased accuracy for quantum many-body problems on quantum computers via an exact transcorrelated method, *Phys. Rev. Res.* **5**, 023174 (2023).
- [108] W. Dobrautz, I. O. Sokolov, K. Liao, P. López Ríos, M. Rahm, A. Alavi, and I. Tavernelli, Ab initio transcorrelated method enabling accurate quantum chemistry on near-term quantum hardware, [arXiv:2303.02007](https://arxiv.org/abs/2303.02007) (2023).

- [109] S. M. Hosseini, A. Alavi, and P. López Ríos, X marks the spot: accurate energies from intersecting extrapolations of continuum quantum Monte Carlo data, Submitted to *J. Chem. Phys.*
- [110] J. Toulouse and C. J. Umrigar, Optimization of quantum Monte Carlo wave functions by energy minimization, *J. Chem. Phys.* **126**, 084102 (2007).
- [111] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, Alleviation of the fermion-sign problem by optimization of many-body wave functions, *Phys. Rev. Lett.* **98**, 110201 (2007).
- [112] C. Filippi and C. J. Umrigar, Multiconfiguration wave functions for quantum Monte Carlo calculations of first-row diatomic molecules, *J. Chem. Phys.* **105**, 213 (1996).
- [113] M. D. Brown, J. R. Trail, P. López Ríos, and R. J. Needs, Energies of the first row atoms from quantum Monte Carlo, *J. Chem. Phys.* **126**, 224110 (2007).
- [114] P. Seth, P. López Ríos, and R. J. Needs, Quantum Monte Carlo study of the first-row atoms and ions, *J. Chem. Phys.* **134**, 084105 (2011).
- [115] F. R. Petruzielo, J. Toulouse, and C. J. Umrigar, Approaching chemical accuracy with quantum Monte Carlo, *J. Chem. Phys.* **136**, 124116 (2012).
- [116] M. A. Morales, J. McMinis, B. K. Clark, J. Kim, and G. E. Scuseria, Multideterminant wave functions in quantum Monte Carlo, *J. Chem. Theory Comput.* **8**, 2181 (2012).
- [117] E. Giner, R. Assaraf, and J. Toulouse, Quantum Monte Carlo with reoptimised perturbatively selected configuration-interaction wave functions, *Mol. Phys.* **114**, 910 (2016).
- [118] A. Scemama, A. Benali, D. Jacquemin, M. Caffarel, and P.-F. Loos, Excitation energies from diffusion Monte Carlo using selected configuration interaction nodes, *J. Chem. Phys.* **149**, 034108 (2018).
- [119] A. Scemama, M. Caffarel, A. Benali, D. Jacquemin, and P.-F. Loos, Influence of pseudopotentials on excitation energies from selected configuration interaction and diffusion Monte Carlo, *Res. Chem.* **1**, 100002 (2019).
- [120] M. Dash, J. Feldt, S. Moroni, A. Scemama, and C. Filippi, Excited states with selected configuration interaction-quantum Monte Carlo: chemically accurate excitation energies and geometries, *J. Chem. Theory Comput.* **15**, 4896 (2019).

- [121] M. Dash, S. Moroni, C. Filippi, and A. Scemama, Tailoring CIPSI expansions for QMC calculations of electronic excitations: the case study of thiophene, *J. Chem. Theory Comput.* **17**, 3426 (2021).
- [122] M. Dash, S. Moroni, A. Scemama, and C. Filippi, Perturbatively selected configuration-interaction wave functions for efficient geometry optimization in quantum Monte Carlo, *J. Chem. Theory Comput.* **14**, 4176 (2018).
- [123] A. Scemama, T. Applencourt, E. Giner, and M. Caffarel, Quantum Monte Carlo with very large multideterminant wavefunctions, *J. Comput. Chem.* **37**, 1866 (2016).
- [124] M. C. Per and D. M. Cleland, Energy-based truncation of multi-determinant wavefunctions in quantum Monte Carlo, *J. Chem. Phys.* **146**, 164101 (2017).
- [125] A. Scemama, Y. Garniron, M. Caffarel, and P.-F. Loos, Deterministic construction of nodal surfaces within quantum Monte Carlo: the case of FeS, *J. Chem. Theory Comput.* **14**, 1395 (2018).
- [126] P. K. V. V. Nukala and P. R. C. Kent, A fast and efficient algorithm for Slater determinant updates in quantum Monte Carlo simulations, *J. Chem. Phys.* **130**, 204105 (2009).
- [127] B. K. Clark, M. A. Morales, J. McMinis, J. Kim, and G. E. Scuseria, Computing the energy of a water molecule using multideterminants: a simple, efficient algorithm, *J. Chem. Phys.* **135**, 244105 (2011).
- [128] G. L. Weerasinghe, P. López Ríos, and R. J. Needs, Compression algorithm for multideterminant wave functions, *Phys. Rev. E* **89**, 023304 (2014).
- [129] C. Filippi, R. Assaraf, and S. Moroni, Simple formalism for efficient derivatives and multi-determinant expansions in quantum Monte Carlo, *J. Chem. Phys.* **144**, 194105 (2016).
- [130] W. Dobrautz, S. D. Smart, and A. Alavi, Efficient formulation of full configuration interaction quantum Monte Carlo in a spin eigenbasis via the graphical unitary group approach, *J. Chem. Phys.* **151**, 094104 (2019).
- [131] M. Caffarel, T. Applencourt, E. Giner, and A. Scemama, Communication: toward an improved control of the fixed-node error in quantum Monte Carlo: the case of the water molecule, *J. Chem. Phys.* **144**, 151103 (2016).

- [132] P. K. Samanta, Private communication.
- [133] P. K. Samanta, D. Mukherjee, M. Hanauer, and A. Köhn, Excited states with internally contracted multireference coupled-cluster linear response theory, *J. Chem. Phys.* **140**, 134108 (2014).
- [134] M. Hanauer and A. Köhn, Pilot applications of internally contracted multireference coupled cluster theory, and how to choose the cluster operator properly, *J. Chem. Phys.* **134**, 204111 (2011).
- [135] J. J. Shepherd, G. H. Booth, A. Grüneis, and A. Alavi, Full configuration interaction perspective on the homogeneous electron gas, *Phys. Rev. B* **85**, 081103(R) (2012).
- [136] J. J. Shepherd, G. H. Booth, and A. Alavi, Investigation of the full configuration interaction quantum Monte Carlo method using homogeneous electron gas models, *J. Chem. Phys.* **136**, 244101 (2012).
- [137] G. G. Spink, R. J. Needs, and N. D. Drummond, Quantum Monte Carlo study of the three-dimensional spin-polarized homogeneous electron gas, *Phys. Rev. B* **88**, 085121 (2013).
- [138] S. Azadi, N. D. Drummond, and S. M. Vinko, Correlation energy of the paramagnetic electron gas at the thermodynamic limit, *Phys. Rev. B* **107**, L121105 (2023).
- [139] Y. Kwon, D. M. Ceperley, and R. M. Martin, Effects of backflow correlation in the three-dimensional electron gas: quantum Monte Carlo study, *Phys. Rev. B* **58**, 6800 (1998).
- [140] M. Holzmann, D. M. Ceperley, C. Pierleoni, and K. Esler, Backflow correlations for the electron gas and metallic hydrogen, *Phys. Rev. E* **68**, 046707 (2003).
- [141] I. G. Gurtubay, R. Gaudoin, and J. M. Pitarke, Benchmark quantum Monte Carlo calculations of the ground-state kinetic, interaction and total energy of the three-dimensional electron gas, *J. Phys.: Condens. Matter* **22**, 065501 (2010).
- [142] S. H. Vosko, L. Wilk, and M. Nusair, Accurate spin-dependent electron liquid correlation energies for local spin density calculations: a critical analysis, *Can. J. Phys.* **58**, 1200 (1980).
- [143] J. P. Perdew and A. Zunger, Self-interaction correction to density-functional approximations for many-electron systems, *Phys. Rev. B* **23**, 5048 (1981).

- [144] J. P. Perdew and Y. Wang, Accurate and simple analytic representation of the electron-gas correlation energy, [Phys. Rev. B **45**, 13244 \(1992\)](#).
- [145] J. Sun, J. P. Perdew, and M. Seidl, Correlation energy of the uniform electron gas from an interpolation between high- and low-density limits, [Phys. Rev. B **81**, 085123 \(2010\)](#).
- [146] M. Holzmann, R. C. Clay, III, M. A. Morales, N. M. Tubman, D. M. Ceperley, and C. Pierleoni, Theory of finite size effects for electronic quantum Monte Carlo calculations of liquids and solids, [Phys. Rev. B **94**, 035126 \(2016\)](#).
- [147] D. Bressanini, G. Morosi, and M. Mella, Robust wave function optimization procedures in quantum Monte Carlo methods, [J. Chem. Phys. **116**, 5345 \(2002\)](#).
- [148] S. M. Hosseini, P. López Ríos, and A. Alavi, Walker number extrapolation of FCIQMC energies, To be published.
- [149] K. Ghanem, A. Y. Lozovoi, and Ali Alavi, Unbiasing the initiator approximation in full configuration interaction quantum Monte Carlo, [J. Chem. Phys. **151**, 224108 \(2019\)](#).
- [150] S. L. Ten-no, Multi-state effective Hamiltonian and size-consistency corrections in stochastic configuration interactions, [J. Chem. Phys. **147**, 244107 \(2017\)](#).
- [151] K. Ghanem, K. Guthrie, and Ali Alavi, The adaptive shift method in full configuration interaction quantum Monte Carlo: development and applications, [J. Chem. Phys. **153**, 224115 \(2020\)](#).
- [152] N. S. Blunt, Communication: an efficient and accurate perturbative correction to initiator full configuration interaction quantum Monte Carlo, [J. Chem. Phys. **148**, 221101 \(2018\)](#).
- [153] B. Ladóczki, M. Uejima, and S. L. Ten-no, Third-order Epstein–Nesbet perturbative correction to the initiator approximation of configuration space quantum Monte Carlo, [J. Chem. Phys. **153**, 114112 \(2020\)](#).
- [154] Q. Sun, X. Zhang, S. Banerjee, P. Bao, M. Barbry, N. S. Blunt, N. A. Bogdanov, G. H. Booth, J. Chen, Z.-H. Cui, J. J. Eriksen, Y. Gao, S. Guo, J. Hermann, M. R. Hermes, K. Koh, P. Koval, S. Lehtola, Z. Li, J. Liu, N. Mardirossian, J. D. McClain, M. Motta, B. Mussard, H. Q. Pham, A. Pulkin, W. Purwanto, P. J. Robinson, E. Ronca, E. R. Sayfutyarova, M. Scheurer, H. F. Schurkus, J. E. T. Smith, C. Sun,

- S.-N. Sun, S. Upadhyay, L. K. Wagner, X. Wang, A. White, J. D. Whitfield, M. J. Williamson, S. Wouters, J. Yang, J. M. Yu, T. Zhu, T. C. Berkelbach, S. Sharma, A. Yu. Sokolov, and G. K.-L. Chan, Recent developments in the P_YSCF program package, *J. Chem. Phys.* **153**, 024109 (2020).
- [155] J. E. T. Smith, B. Mussard, A. A. Holmes, and S. Sharma, Cheap and near exact CASSCF with large active spaces, *J. Chem. Theory Comput.* **13**, 5468 (2017).
- [156] A. A. Holmes, C. J. Umrigar, and S. Sharma, Excited states using semistochastic heat-bath configuration interaction, *J. Chem. Phys.* **147**, 164111 (2017).
- [157] S. M. Casey and D. G. Leopold, Negative ion photoelectron spectroscopy of chromium dimer, *J. Phys. Chem.* **97**, 816 (1993).
- [158] N. B. Balabanov and K. A. Peterson, Systematically convergent basis sets for transition metals. I. All-electron correlation consistent basis sets for the 3*d* elements Sc–Zn, *J. Chem. Phys.* **123**, 064107 (2005).
- [159] W. Kutzelnigg and W. Liu, Quasirelativistic theory equivalent to fully relativistic theory, *J. Chem. Phys.* **123**, 241102 (2005).
- [160] M. Douglas and N. M. Kroll, Quantum electrodynamical corrections to the fine structure of helium, *Ann. Phys.* **82**, 89 (1974).
- [161] B. A. Hess, Relativistic electronic-structure calculations employing a two-component no-pair formalism with external-field projection operators, *Phys. Rev. A* **33**, 3742 (1986).