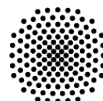# Development of the Transcorrelated Full Configuration Interaction Quantum Monte Carlo Method

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

Vorgelegt von
Jacobus Philip Haupt
aus George, Südafrika

Hauptberichter:　　　　　　Prof. Dr. Ali Alavi
Mitberichter:　　　　　　　Prof. Dr. Andreas Köhn
Prüfungsvorsitzender:　　Prof. Dr. Joris van Slageren
Tag der mündlichen Prüfung:　30 April 2025

Universität Stuttgart

Max-Planck-Institut für Festkörperforschung, Stuttgart

2025

It is important to realize that in physics today, we have no knowledge of what energy is.

— Richard Feynman, *Feynman Lectures on Physics, Volume 1, Chapter 4*

CONTENTS

Erklärung über die Eigenständigkeit der Dissertation

Ich erkläre hiermit an Eides statt, dass ich die vorliegende Arbeit

„Development of the Transcorrelated Full Configuration Interaction Quantum Monte Carlo Method "

ohne Hilfe Dritter und ohne Benutzung anderer als der angegebenen Hilfsmittel angefertigt habe; die aus fremden Quellen direkt oder indirekt übernommenen Gedanken sind als solche kenntlich gemacht. Die Arbeit wurde bisher in gleicher oder ähnlicher Form in keiner anderen Prüfungsbehörde vorgelegt.

_____

Datum, Ort   Unterschrift

Acknowledgements

Foremost, I would like to thank my supervisor **Ali Alavi** for his guidance and undying optimism, as well as for being an excellent professional role model. Similarly, I would like to thank **Daniel Kats** for his professional guidance and patience with me as I learned the basics of electronic structure theory, especially at the start of my PhD while I was still working remotely from Vancouver. **Nikolay Bogdanov** has also been extremely helpful in guiding me with active-space methods. I also extend a thanks to my committee members **Joris van Slageren** and **Andreas Köhn**.

Similarly, much of my professional development can be owed to mentors along the way, in particular **Darryl Barber**, **Christian Ast**, and **Roman Krems**.

I also thank **Evelin Christlmaier** for all her support, both professionally and personally, and with whom working has been a pleasure.While our time together in the group was short, **Werner Dobrautz** distilled in me a greater sense of mindfulness and reassurance that everything will be okay. Similarly, **Johannes Hauskrecht** has been a wonderful officemate with whom I could always whine about life as a PhD student.

My time at MPI has, to my pleasure, involved a lot of software development, and for that I am especially grateful to have had such knowledgeable coworkers as **Robert Anderson** and **Oskar Weser** to go to for advice and with whom to argue which text editor is best.

I would also like to thank my former and present coworkers **Pablo López Ríos**, **Aron Cohen** and **Kai Guther** for their work on making `tchint` a reality. **Andreea Filip**, **Kristoffer Simula** and **Ke Liao** have been especially pivotal in making the `pytchint` package possible.

Navigating administration as a foreigner in Germany can be extremely difficult, but this stress has been lifted off my shoulders by **Birgit King**, to whom I am much in debt.

Finally, I want to thank my dear friends who made my time in Stuttgart infinitely more colourful: **Aybike Reyhanlı**, **Thomas Schraivogel**, **Ingemar Schmidt**, **Rut Martinez**, **Nikos Papanikolaou**, **Aimée Sousa Calepso**, **Martin Baumann**, and especially **Theresa Huber** with whom I shall always cherish our countless hours dancing together.

Acronyms

SD Slater determinant 7, 11, 15, 16, 19, 25, 29, 42–44, 49

STO Slater-type orbital 8, 9, 22, 103

TC transcorrelation 2, 28, 49–51, 86, 87, 101, 141, 143

UHF unrestricted Hartree-Fock 12

VMC variational Monte Carlo 29, 37, 38, 41, 53, 54, 101, 144

xTC transcorrelation via exclusion of explicit three-body components 68, 110, 142, 143, 146

Zusammenfassung

Die Transkorrelierte (TC) Methode ist eine Methode in der Quantenchemie, die in den letzten Jahren an Bedeutung gewonnen hat. In dieser Methode wird eine Ähnlichkeitstransformation auf den elektronischen Hamiltonoperator angewandt, um Korrelationseffekte, insbesondere die dynamische Korrelation, durch explizite Behandlung analytisch bekannter Eigenschaften des Hamiltonoperators zu erfassen. Diese Methode wurde bereits mit der „Full Configuration Interaction Quantum Monte Carlo" (FCIQMC) Methode kombiniert, einer effizienten stochastischen Methode zur Lösung der Schrödingergleichung und zur Berechnung physikalischer Observablen. Diese Dissertation widmet sich einer vertieften Untersuchung dieser Methoden, insbesondere der TC-Methode.

Nach einer Wiederholung der Grundlagen wellenfunktionsbasierter und Quantum-Monte-Carlo Methoden, wird die Verwendung flexibler Jastrow-Faktoren, die in der Literatur des „Variational Quantum Monte Carlos" (VMC) bekannt sind, vorgestellt. Damit minimieren wir die Varianz der TC-Referenzenergie. Es zeigt sich, dass dies zu einer schnellen Basissatzkonvergenz führt und eine Genauigkeit erreicht, für die herkömmliches FCIQMC viel größere Basissätze erfordern würde. Außerdem kompaktiert dieses Minimierungsverfahren auch die Wellenfunktion, was eine effizientere FCIQMC-Berechnung ermöglicht. Im weiteren Verlauf erweitern wir die Methode für Probleme mit starken Multireferenzcharakter, insbesondere indem wir die Dissoziation des Stickstoffdimers als Stresstest verwenden. Wir zeigen die Notwendigkeit eines Multireferenz-Jastrow-Faktor-Ansatzes auf und minimieren deshalb die Varianz eines Multireferenzzustands. Dies wird nachweislich günstige, größenkonsistente Energien wiederherstellen und gleichzeitig die schnelle Basissatzkonvergenz von TC erreichen. Da nach der Jastrow-Faktor-Optimierung eine Multireferenztechnik (FCIQMC) verwendet wird, erhöht sich der rechnerische Skalierung für die Methode darüber hinaus nicht, wenn man die konventionelle (nicht-TC-)Form derselben Multireferenztechnik wie beim TC-Ansatz verwendet.

Abschließend untersuchen wir die Möglichkeit, vereinfachte Jastrow-Faktoren zu konstruieren, um das bisherige Optimierungsverfahren für die TC-Methode, das rechenintensiv sein kann, zu verbessern oder sogar ganz zu umgehen. Einerseits zeigen wir, dass parameterfreie Jastrow-Faktoren aufgrund der Fehlerkompensation zu schlechten absoluten Energien, aber guten relativen Energien führen können. Andererseits ergeben Jastrow-Faktoren, die für Atome optimiert und für Moleküle wiederverwendet wurden, sowohl genaue absolute als auch relative Energien. Dies eröffnet die Möglichkeit, Jastrow-Faktoren für Atome im gesamten Periodensystem zu optimieren und sie in einer Datenbank zu

speichern, die für größere Moleküle abgefragt werden kann, was die Skalierbarkeit der Methode unterstützt.

Abstract

The transcorrelated (TC) method is a technique in electronic structure theory that has recently been gaining momentum. In it, a similarity transformation is applied to the electronic Hamiltonian to capture effects of electron correlation, particularly dynamical correlation, by explicit treatment of analytically-known properties of the Hamiltonian near coalescence points. This has already been combined with the full configuration interaction quantum Monte Carlo (FCIQMC) method, an efficient stochastic approach to solving the electronic Schrödinger equation and calculating physical observables. This dissertation further explores these methods, particularly TC.

After a recapitulation of the core concepts of wave function and quantum Monte Carlo methods, we introduce the use of flexible Jastrow factors familiar in the variational Monte Carlo (VMC) literature to minimise the variance of the TC-reference energy. This is shown to result in a rapid basis-set convergence, reaching accuracy for which conventional FCIQMC would require much larger basis sets. Moreover, this minimisation procedure is shown to also compactify the wave function, allowing for more efficient sampling in FCIQMC. We next extend the methodology for problems of strongly multireference character, notably using the dissociation of the nitrogen dimer as a stress test. We illustrate the need for a multireference Jastrow-factor ansatz, and hence minimise the variance of a multireference state. This is shown to recover favourable, size-consistent energies while maintaining the rapid basis-set convergence that comes with TC. As a multireference technique (FCIQMC) is used after optimisation, it moreover does not increase the computational scaling of the method to use the conventional (non-TC) form of the same multireference technique as the TC ansatz.

Finally, we explore the possibility of constructing simplified Jastrow factors in order to improve or possibly wholly bypass the optimisation procedure so far for the TC method, which can be computationally expensive. We show that parameter-free Jastrow factors can result in poor absolute energies, but favourable relative energies thanks to error cancellation, whereas Jastrow factors optimised for atoms and reused for molecules result in both accurate absolute and relative energies. This opens the possibility of optimising Jastrow factors for atoms across the periodic table and storing them in a database, which can be queried for larger molecules, thereby aiding the scalability of the method.

Introduction

According to modern quantum theory, to fully describe a (nonrelativistic) system of N_A nuclei and N_e electrons, we must solve the time-dependent Schrödinger equation,

$$i\hbar \frac{\partial}{\partial t}\Psi(\boldsymbol{x}, t) = \hat{H}\Psi(\boldsymbol{x}, t) \tag{1.1}$$

where $\boldsymbol{x} = (\boldsymbol{r}, \sigma)$, $\boldsymbol{r} \in \mathcal{R}^{3(N_e+N_A)}$ gives the spatial coordinates, and σ the corresponding spins. For nuclei with atomic numbers Z_I (positive charge $Z_I e$), the Hamiltonian operator $\hat{H}$ may be written

$$
\begin{aligned}
\hat{H} = &-\sum_{I=1}^{N_A} \frac{\hbar^2}{2m_I}\nabla_I^2 + \frac{1}{4\pi\epsilon_0}\frac{1}{2}\sum_{I,J=1;I\neq J}^{N_A} \frac{Z_I Z_J e^2}{|\boldsymbol{R}_I - \boldsymbol{R}_J|} \\
&-\sum_{i=1}^{N_e} \frac{\hbar^2}{2m_e}\nabla_i^2 + \frac{1}{4\pi\epsilon_0}\frac{1}{2}\sum_{i,j=1;i\neq j}^{N_e} \frac{e^2}{|\boldsymbol{r}_i - \boldsymbol{r}_j|} - \frac{1}{4\pi\epsilon_0}\sum_{i=1}^{N_e}\sum_{I=1}^{N_A} \frac{Z_I e^2}{|\boldsymbol{r}_i - \boldsymbol{R}_I|}
\end{aligned}
\tag{1.2}
$$

where N_A and N_e are the number of nuclei and electrons, respectively, and we have represented the nuclear coordinates by $\boldsymbol{R}_I$ and the electron coordinates by $\boldsymbol{r}_i$.

Since quantum chemistry and condensed matter sciences are in general concerned with nonrelativistic processes involving electrons and nuclei, this might boldly be called the *theory of everything.*[1] Hence, we may be tempted to conclude this dissertation early, but in practice the evaluation of the Hamiltonian in equation 1.2 is impossible. First, there is no closed form solution of equation 1.1 with this Hamiltonian. Second, numerical evaluation of $\hat{H}$ is far from trivial.

Let's consider a simple example, the Argon atom. Say we want to solve this partial differential equation on a grid. Let's choose a very coarse $10 \times 10 \times 10$ grid. Then *at each time step* we need to store $10^{3 \times (18+18)} = 10^{108}$ values, corresponding to all the particle positions and the grid. Considering there are "only" $\sim 10^{80}$ atoms in the known universe,[2] this is completely unreasonable.

This system has only 18 electrons and 18 nuclei, a far cry from the 10^{23} or higher number of electrons in a typical condensed matter system, for example. Moreover, we have not taken into account floating point precision, or that we would need to calculate

this for possibly many time steps. Clearly, drastic approximations and more sophisticated methods are required.

1.1. Overview of the Thesis

This dissertation fits into the field of nonrelativistic electronic structure theory, the branch of quantum chemistry concerned with the description of electrons and their correlation inside molecules and materials. More specifically, this dissertation focuses on high-accuracy (but often high-cost) *ab initio* methodologies, especially full configuration interaction quantum Monte Carlo (FCIQMC) and transcorrelation (TC). As such, we will only be discussing small systems consisting of only a few atoms, as they are tractible with a full, all-electron treatment with these methods. In principle, these methods should be able to be embedded[3,4] into more large-scale calculations using multiscale techniques, but this is outside the scope of my work. Nevertheless, the work herein is focused on methodologies, and not on particular physical systems.

The outline of the dissertation is as thus:

- Chapter 1 (this chapter) provides a basic overview of electronic structure theory methods and some of its principal concepts. Sections 1.5 and 1.6 in particular are largely based on the appropriate chapters of reference 5.

- Chapter 2 reviews the current works in so-called "explicitly correlated" methods, notably the well-established R12/F12 and the recently-reinvigorated TC.

- Chapter 3 provides a basic introduction to quantum Monte Carlo (QMC) and how it relates to FCIQMC.

- Chapter 4 discusses optimization strategies of Jastrow factors in the context of TC.

- Chapter 5 discusses an extension of the methods in the previous chapters to ensure size consistency and success when targeting strongly multireference problems.

- Chapter 6 discusses methods for modular and simple Jastrow factor forms, for easier scalability.

- The final chapter, chapter 7, provides a review and an outlook for the field.

- Appendix A provides an overview of the software `pytchint` developed in the group for evaluation of TC integrals.

1.2. Principal Approximations

As discussed, to make any progress in electronic structure theory, we must make use of approximations. Of course, we must always be cautious and suspicious when using these, and make sure they are valid for the systems in question. Thankfully, over the course of the last century, a cornucopia of different approximations has been developed to tackle the Schrödinger question, some of which will be discussed in this section.

1.2.1. The Born-Oppenheimer Approximation

The most important approximation used in electronic structure theory is the Born-Oppenheimer approximation (BOA).[6] The BOA relies on the fact that the nuclei are much heavier than the electrons, with the mass of a single proton being almost 2000 times the mass of an electron. As an intuitive picture, we may think of the nuclei as moving much slower than the electrons, which can adapt themselves to the instantaneous positions of the nuclei. In mathematical terms, this means we can take the total wave function to be a product of its nuclear and electronic components,

$$\Psi_{\text{total}} = \Psi_{\text{nuc}}\Psi_{\text{elec}}. \tag{1.3}$$

Notice that the first two terms of equation 1.2 are independent of the electronic coordinates and, ipso facto, have no effect on Ψ_{elec}. This leads to the *electronic Hamiltonian* under the BOA, which can be written as

$$\hat{H}_{\text{elec}} = -\sum_i \frac{1}{2}\nabla_i^2 - \sum_{i,I} \frac{Z_I}{r_{iI}} + \sum_{i>j} \frac{1}{r_{ij}}, \tag{1.4}$$

where we have simplified notation by using miniscule roman letters for the electrons and capital roman letters for the nuclei, and $r_{ij} = |\mathbf{r}_i - \mathbf{r}_j|$ and $r_{iI} = |\mathbf{r}_i - \mathbf{R}_I|$, as well as by using atomic units. That is, the system of units where $\hbar$, e, m_e, and $4\pi\epsilon_0$ all correspond to the value of 1.

In the language of second quantisation, equation 1.4 can be written as

$$\hat{H}_{\text{elec}} = \sum_\sigma \sum_{pq} h_q^p a_{p\sigma}^\dagger a_{q\sigma} + \frac{1}{2} \sum_{\sigma\tau} \sum_{pqrs} V_{rs}^{pq} a_{p\sigma}^\dagger a_{r\tau}^\dagger a_{s\tau} a_{q\sigma}, \tag{1.5}$$

where p, q, r, s are general spatial-orbital indices, σ, τ are the spin indices, and $a_{p\sigma}$ $(a_{p\sigma}^\dagger)$ is the annihilation (creation) operator for an electron in spatial-orbital p with spin σ. These must obey the anti-commutation relations

$$\{a_{p\sigma}, a_{q\tau}^\dagger\} := a_{p\sigma}a_{q\tau}^\dagger + a_{q\tau}^\dagger a_{p\sigma} = \delta_{pq}\delta_{\sigma\tau} \tag{1.6}$$

$$\{a_{p\sigma}, a_{q\tau}\} = \{a_{p\sigma}^\dagger, a_{q\tau}^\dagger\} = 0 \tag{1.7}$$

so that the electrons satisfy the Pauli exclusion principle.

In equation 1.5, h_q^p is a one-body integral,

$$h_q^p = \int \mathrm{d}^3 r \; \phi_p^*(\mathbf{r}) \left(-\frac{1}{2}\nabla^2 - \sum_I \frac{Z_I}{|\mathbf{R}_I - \mathbf{r}|} \right) \phi_q(\mathbf{r}) \tag{1.8}$$

and V_{rs}^{pq} is a two-body integral,

$$V_{rs}^{pq} = \int \mathrm{d}^3 r \int \mathrm{d}^3 r' \; \phi_p^*(\mathbf{r})\phi_r^*(\mathbf{r}') \frac{1}{|\mathbf{r} - \mathbf{r}'|} \phi_q(\mathbf{r})\phi_s(\mathbf{r}'), \tag{1.9}$$

with a spatial-orbital basis $\{\phi_i(\mathbf{r})\}$.

For our purposes, we neglect the first term of equation 1.2, and treat the second term as approximately constant, so $\hat{H}_{\mathrm{nuc}} = \sum_{IJ} Z_I Z_J r_{IJ}^{-1}$. Thus, we are left with the problem of solving the electronic structure problem, which is the subject of this dissertation. Solving the electronic problem for different nuclear positions gives us the potential energy surface.

Note that while chemists and physicists typically talk about the positions of nuclei in a molecule or solid as if they are fixed in place, as will be done in this dissertation, this is really a colloquialism. If the nuclei had an exact position and zero kinetic energy, the BOA would be in direct contradiction of the Heisenberg uncertainty principle. Instead, on the timescale of the electrons, due to the much higher mass of nuclei, in the BOA we treat the nuclei as approximately localised in a state in which their motion is much slower than that of the electrons (but, importantly, not zero). This keeps the approximation from being in conflict with the fundamental postulates of quantum theory.

The BOA is an immensely practical tool as it substantially simplifies our equations, and in many applications it is an excellent approximation. It will be a fundamental assumption throughout the rest of this dissertation, though it need not always be valid in all of quantum chemistry.

While we have done a lot to drastically reduce the complexity of equation 1.2, equation 1.5 is still intractible for large system sizes, scaling combinatorially with the size of the Hilbert space, as a function of the system size N_e (henceforth N), and the basis set size M. Hence, in addition to using a smaller basis, we still need extra approximations and sophisticated methodologies.

1.2.2. Core Orbitals

In the ground state of a single atom, electrons first occupy the lowest orbitals. The first occupied shells are typically where the electrons are the most tightly bound.*

Since they tend to be further from the nucleus in an atom, the valence (non-core) orbitals are typically the most affected by the introduction of additional atoms in the

---

*This neglects symmetries: energy may be higher due to the angular momentum while staying spatially closer to the nucleus.

system. It is therefore often the valence orbitals that are most important in chemical systems. For this reason, we sometimes "freeze" the core orbitals, meaning they remain doubly occupied by electrons, and only correlate the valence orbitals. It is also common to refer to core electrons and valence electrons, meaning those electrons occupying the core and valence orbital space, respectively (though this is technically imprecise language, as electrons are indistinguishable).

Similarly, we may also delete virtual orbitals, typically those of high energy, to further reduce the size of the problem. The remaining space is known as an active space, and is the basic idea of the complete active space (CAS) methods.

1.2.3. Model Hamiltonians

In 1929, the surrealist artist René Magritte displayed a now-famous painting of a pipe with the caption *Ceci n'est pas une pipe* (French for "This is not a pipe"). It was meant to depict the idea that the painting itself is in a way treacherous: it may appear to be a pipe, but you cannot stuff it or smoke from it, as it is a representation of a pipe.

In a similar way, physicists and chemists often use *model Hamiltonians*, which do away with aspects of equation 1.2 that are not expected to be relevant to the problem at hand, resulting in new Hamiltonians that may be considered a representation of equation 1.2, much like Magritte's painting.* Compared to *ab initio* Hamiltonians like equation 1.5, model Hamiltonians are generally much simpler, but depend on parameters whose values we may not necessarily know a priori.

The most famous model Hamiltonian is the Hubbard model,[7] most typically used to describe electrons in a periodic lattice, and can be written as

$$\hat{H}_{\text{Hubb}} = -\sum_{\langle p,q \rangle} \sum_{\sigma} t_{pq} a_{q\sigma}^{\dagger} a_{p\sigma} + \sum_{p} U_{p} \hat{n}_{p\uparrow} \hat{n}_{q\downarrow}, \tag{1.10}$$

where $\langle p,q \rangle$ denotes nearest neighbour sites, t_{pq} (often taken to be constant, t) is a parameter called the hopping amplitude, and U_p (often taken to be constant, U) is a parameter called the on-site repulsion. Here we have used spatial-orbital basis labels $\{p,q\}$ and spin labels $\sigma \in \{\uparrow,\downarrow\}$. $\hat{n}_{p\sigma} := a_{p\sigma}^{\dagger} a_{p\sigma}$ is the number operator.

Despite its apparent simplicity, the Hubbard model is a rich model, and has been used to describe a variety of phenomena, from metal-insulator transitions to high-temperature superconductivity. Moreover, it is not always so simple to solve, and has been the subject of much research.[8,9]

---

*We might further argue that all of science can be described this way, as scientific models are always "mere" representations of Nature, and are not Nature itself.

Let's consider an especially simple version of the model, with only two sites, and open boundary conditions:

$$\hat{H}_{\text{Hubb}} = -t \sum_\sigma (a_{1\sigma}^\dagger a_{2\sigma} + a_{2\sigma}^\dagger a_{1\sigma}) + U(\hat{n}_{1\uparrow}\hat{n}_{1\downarrow} + \hat{n}_{2\uparrow}\hat{n}_{2\downarrow}). \qquad (1.11)$$

We could use this as an approximate description of the H_2 molecule, where the electrons are restricted to only the lowest 1s orbitals. Compared to equation 1.2, this is a tremendously easy problem now.

Of course, when solving a problem using a model Hamiltonian, we must be careful in choosing the correct one. In particular, the exact solution to the Hubbard model in one dimension[8] does not appear to be in agreement with the *ab initio* solution of a one-dimensional chain of hydrogen atoms.[10,11]

1.3. The Variational Principle

One way of finding approximations to the lowest energy eigenstate of a quantum problem like that of equation 1.5 is the variational method. Several methods discussed in this dissertation build on the variational method, and the variational method itself is based on the variational principle, which states[12] for any "trial wave function" $\left|\tilde{\Psi}\right\rangle$ and Hermitian operator (such as the electronic Hamiltonian) $\hat{H}$,

$$\frac{\left\langle \tilde{\Psi}\right| \hat{H} \left|\tilde{\Psi}\right\rangle}{\left\langle \tilde{\Psi}\middle|\tilde{\Psi}\right\rangle} \geq E_0 := \frac{\langle\Psi| \hat{H} |\Psi\rangle}{\langle\Psi|\Psi\rangle} \qquad (1.12)$$

where E_0 is the lowest eigenvalue (e.g. the ground state energy). That is, the energy of $\left|\tilde{\Psi}\right\rangle$ is an upper bound of the exact energy of the system.

To prove this, consider an expansion of $\left|\tilde{\Psi}\right\rangle$ terms of the exact eigenkets $|k\rangle$ of $\hat{H}$, so that $\hat{H}|k\rangle = E_k |k\rangle$ and $\left|\tilde{\Psi}\right\rangle = \sum_k c_k |k\rangle$. Then

$$\begin{aligned}
\frac{\left\langle \tilde{\Psi}\right| \hat{H} \left|\tilde{\Psi}\right\rangle}{\left\langle \tilde{\Psi}\middle|\tilde{\Psi}\right\rangle} &= \frac{\sum_{kk'} c_k^* c_{k'} \langle k| \hat{H} |k'\rangle}{\sum_{kk'} \langle k|k'\rangle c_k^* c_{k'}} \\
&= \frac{\sum_k |c_k|^2 (E_k - E_0)}{\sum_k |c_k|^2} + E_0 \\
&\geq E_0
\end{aligned} \qquad (1.13)$$

where, in going from the first to the second line, we used the orthonormality of the eigenbasis $\{|k\rangle\}$ and that the left- and right-eigenvectors are the same. That is, we used the fact the $\hat{H}$ is a Hermitian operator. This is an important observation for future discussions, such as in section 2.5.

1.4. The Schrödinger Equation as a Matrix Problem

Implicit in the operator algebra of the fermionic creation and annihilation operators $a^\dagger$ and a in equation 1.5 is the Pauli exclusion principle: we cannot have two or more fermions occupying identical states. In the context of electronic structure theory and this dissertation, the fermions and states in question are electrons and orbitals, respectively. This handles the antisymmetry requirement of the fermionic wave function, $\hat{P}_{ij}^\sigma \Psi = -\Psi$, where $\hat{P}_{ij}^\sigma$ permutes the ith and jth electrons with spin σ.

Our wave function can hence be expanded by a linear combination of Slater determinants (SDs),

$$\Psi = \sum_i c_i \left| D_i \right\rangle, \quad \text{where } \left| D_i \right\rangle = \frac{1}{\sqrt{N!}} \begin{vmatrix} \chi_1(\boldsymbol{x}_1) & \chi_1(\boldsymbol{x}_2) & \cdots & \chi_1(\boldsymbol{x}_N) \\ \chi_2(\boldsymbol{x}_1) & \chi_2(\boldsymbol{x}_2) & \cdots & \chi_2(\boldsymbol{x}_N) \\ \vdots & \vdots & \ddots & \vdots \\ \chi_N(\boldsymbol{x}_1) & \chi_N(\boldsymbol{x}_2) & \cdots & \chi_N(\boldsymbol{x}_N) \end{vmatrix} \tag{1.14}$$

for a spin-orbital basis $\{\chi_p(\boldsymbol{x}_q)\}$. A spin-orbital is typically written as a product of a spatial part and a spin part, e.g. $\chi(\boldsymbol{x}) = \phi(\boldsymbol{r})\omega(\sigma)$.

Given such a finite spin-orbital basis of size $2M$ (M spatial orbitals with 2 different spins each),[*] the SD expansion is also finite. Therefore, we may write $\hat{H}$ as a square matrix and equation 1.5 as an eigenvalue equation,

$$\hat{H}\Psi = E\Psi. \tag{1.15}$$

Note, however, that in order to make this a finite-dimensional matrix eigenvalue problem, we needed to use a finite spin-orbital basis. Thus, E in equation 1.15 is not the true ground state energy of the system, but indeed only an approximation within that basis set. This is the case even if we were to solve equation 1.15 exactly. In order to get the true ground state energy, we need to reach the complete basis set (CBS), i.e. $M \to \infty$.

With equation 1.15, we have transformed the hopelessly intractable partial differential equation 1.2 into a finite algebraic eigenvalue problem, which is much better suited for solving on a computer.

If we have $N_\uparrow$ spin-up electrons and $N_\downarrow$ spin-down electrons, we then have

$$|\mathcal{H}| = \binom{M}{N_\uparrow}\binom{M}{N_\downarrow} \tag{1.16}$$

where $|\mathcal{H}|$ denotes the size of the Hilbert space, and we have assumed that the number of spin-up orbitals M is equal to the number of spin-down orbitals. Here, $\binom{n}{r} = \frac{n!}{r!(n-r)!}$ is a binomial coefficient.

---

[*]We assume here the same number of spin-↑ and spin-↓ orbitals for simplicity, though it need not always be the case.

The combinatorial scaling is still unfavourable. For a closed-shell system with 20 electrons in 20 orbitals, $|\mathcal{H}| = 3.41 \times 10^{10}$. Assuming we wish to have double precision,* to store the whole matrix we would need about 10^{23} bytes, or 100 zettabytes. So, while we've substantially reduced the complexity of the problem, it is still extremely demanding. More approximations are therefore needed.

1.4.1. Orbitals

Orbitals (and spin-orbitals), a term already used in this text, describe the spatial distribution (and in the case of spin-orbitals, also the spin) of a single electron. We assume that spatial orbitals form an orthonormal set, and that they form a complete space for an arbitrary wave function. However, in theory to have a CBS would mean needing an infinite number of basis functions. In practice, we must of course have a finite basis, and so this naturally leads to a hierarchy of basis sets, where the typically larger basis sets are both more expensive and more accurate.

Many different basis set families exist, but among the most popular are those due to Dunning and coworkers.[14] Additionally, many different basis set extrapolation techniques have also been developed.[15–22]

Strictly speaking, any kind of set of functions can be used as a basis set, as long as they are a complete set. However, in practice we may broadly categorise them into two groups: Slater-type orbitals (STOs) and Gaussian-type orbitals (GTOs).

The exact solution for the Hydrogen atom is the Slater-type function $\frac{1}{\sqrt{\pi}}\mathrm{e}^{-r}$, and similarly we know that molecular orbitals also decay as $\sim e^{-\zeta r}$.

However, these STOs are difficult to work with, as integrals of the form

$$\int \mathrm{d}^3 r_1 \int d^3 r_2 \ \phi_p^*(\boldsymbol{r}_1)\phi_r^*(\boldsymbol{r}_2)\frac{1}{r_{12}}\phi_q(\boldsymbol{r}_1)\phi_s(\boldsymbol{r}_2) \tag{1.17}$$

are routinely needed.

Therefore, GTOs have been developed,[23] which have the form $\mathrm{e}^{-\alpha r^2}$. GTOs are much easier to handle, as the product of a gaussian is also a gaussian, and the integral of a gaussian is exactly known, so equations like 1.17 have known solutions. With this comes a tradeoff, since we know STOs more faithfully capture the form of molecular orbitals, but GTOs are much better suited for computations. For this, the most common basis sets typically use a linear combination (or "contraction") of gaussians fitted to a STO. An illustration is shown in figure 1.1.

It must be noted, however, that STOs and GTOs have different asymptotic behaviour for $r \to \infty$ and $r \to 0$. The latter in particular will be discussed in section 2.1.

---

*A double precision floating-point number has 1 bit for the sign, 11 bits for the exponent, and 52 bits for digits in the number (known as the mantissa or significand).[13] This resolves to 64 bits or 8 bytes.
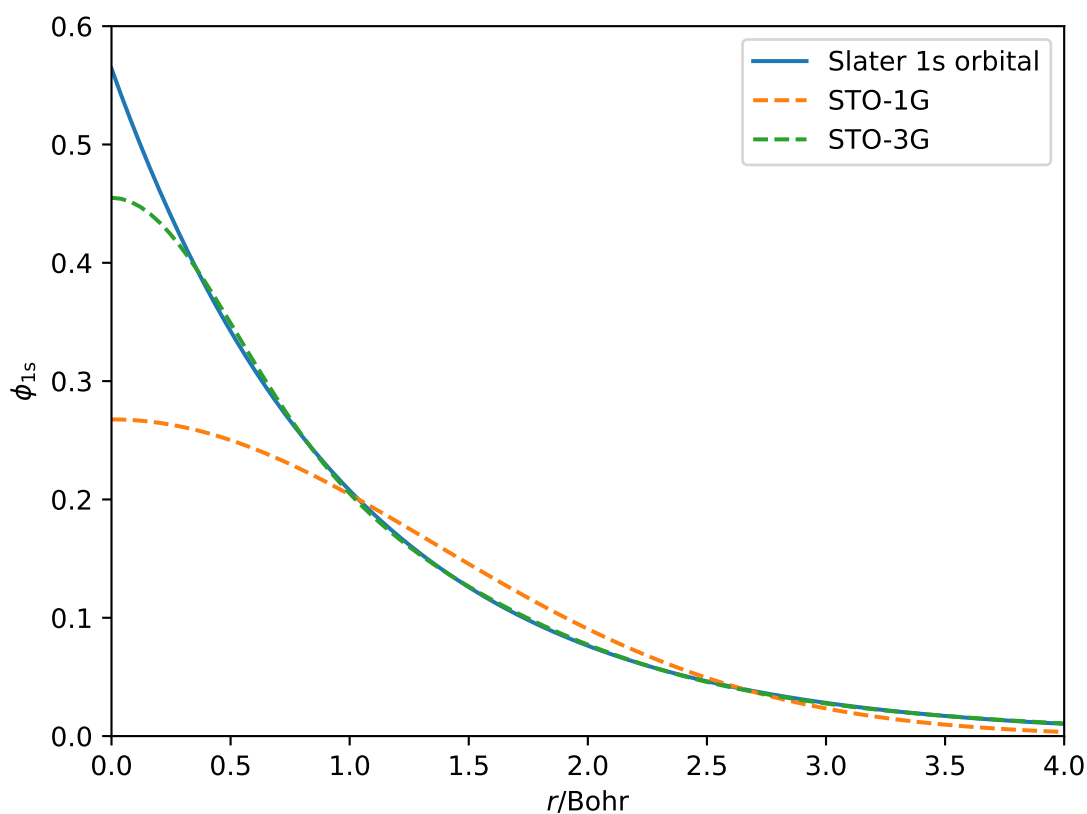
Figure 1.1. Illustration of an s-type orbital and the fitted gaussians (STO-xG means x gaussians are used to approximate the STO). Note that for much of the curve, there is good agreement, but not for the short- and long-range. The short-range behaviour in particular leads to complications, and is the focus of the methods discussed in chapter 2. The long-range behaviour is typically less of a problem, since the integrals we are interested in typically decay with r. Fitting parameters for the gaussians were taken from reference 24.

1.4.2. Density Matrices

The 1-electron reduced density matrix (1RDM) is defined by its matrix elements

$$\gamma_{pq} = \langle \Psi | \, a_q^\dagger a_p \, | \Psi \rangle, \tag{1.18}$$

or written in a spin-orbital basis:

$$\gamma(\boldsymbol{x}_1, \boldsymbol{x}_2) = \sum_{pq} \phi_q(\boldsymbol{x}_2)^* \phi_p(\boldsymbol{x}_1) \gamma_{pq} \tag{1.19}$$

Similarly, the 2-electron reduced density matrix (2RDM) is defined as

$$\Gamma_{pqrs} = \langle \Psi | \, a_s^\dagger a_r^\dagger a_q a_p \, | \Psi \rangle. \tag{1.20}$$

Any one- and two-electron Hermitian operator Ω can be written (in a spin-orbital basis) as

$$\langle \Psi | \, \Omega \, | \Psi \rangle = \sum_{pq} \gamma_{pq} \Omega_{pq} + \sum_{pqrs} \Gamma_{pqrs} \Omega_{rpsq}. \tag{1.21}$$

1.4.3. Electron Correlation

Two variables are independent (or uncorrelated) if the joint probability distribution is the product of their expected values, i.e.

$$P(\boldsymbol{x}_1, \boldsymbol{x}_2) = P(\boldsymbol{x}_1) P(\boldsymbol{x}_2). \tag{1.22}$$

By Bayes' theorem,[25,26] this can be rewritten in terms of the conditional probability,

$$P(\boldsymbol{x}_1 | \boldsymbol{x}_2) = \frac{P(\boldsymbol{x}_1, \boldsymbol{x}_2)}{P(\boldsymbol{x}_2)} = \frac{P(\boldsymbol{x}_1) P(\boldsymbol{x}_2)}{P(\boldsymbol{x}_2)} = P(\boldsymbol{x}_1). \tag{1.23}$$

The variables are said to be correlated if the above is not true.

In the case of electronic structure, the variables in question are $\boldsymbol{x} := (\boldsymbol{r}, \sigma)$, the spatial coordinates and spin of the electrons. Therefore, when we speak of electron correlation, we are referring to these relations. Furthermore, as electrons are indistinguishable, for every pair $(\boldsymbol{x}_1, \boldsymbol{x}_2)$,

$$P(\boldsymbol{x}_1) = P(\boldsymbol{x}_2) = \frac{1}{N} \rho(\boldsymbol{x}) \tag{1.24}$$

$$P(\boldsymbol{x}_1, \boldsymbol{x}_2) = \frac{1}{N(N-1)} \rho(\boldsymbol{x}_1, \boldsymbol{x}_2) \tag{1.25}$$

where $\rho(\boldsymbol{x})$ is the electron density and $\rho(\boldsymbol{x}_1, \boldsymbol{x}_2)$ is the probability of finding electrons at $(\boldsymbol{x}_1, \boldsymbol{x}_2)$ simultaneously, known as the pair density,

$$\rho(\boldsymbol{x}) = N \int \mathrm{d}^3 x_2 \cdots \mathrm{d}^3 x_N \Psi^*(\boldsymbol{x}, \boldsymbol{x}_2, \boldsymbol{x}_3, \cdots, \boldsymbol{x}_N) \Psi(\boldsymbol{x}, \boldsymbol{x}_2, \boldsymbol{x}_3, \cdots, \boldsymbol{x}_N) \tag{1.26}$$

$$\rho(\boldsymbol{x}_1, \boldsymbol{x}_2) = N(N-1) \int \mathrm{d}^3 x_2 \cdots \mathrm{d}^3 x_N \Psi^*(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \cdots, \boldsymbol{x}_N) \Psi(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{x}_3, \cdots, \boldsymbol{x}_N). \tag{1.27}$$

There are two key sources of electron correlation: Fermi correlation, which arises from the fact that electrons obey Fermi statistics, i.e. the wave function is antisymmetric with respect to exchange of $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$, and Coulomb correlation, which arises from the fact that electrons repel each other.

1.5. The Hartree-Fock Method

The typical starting point to most *ab initio* electronic structure methods is the Hartree-Fock (HF) method.[27–29] The key approximation in HF is that we treat the exact N-body wave function solution as approximately a single SD.* Then, we variationally (see section 1.3) optimise the energy of the system with respect to parameters pertaining to the spin-orbitals. We will refer to this SD as $|\Phi_0\rangle$.

Then, for some arbitrary initial configuration $|0\rangle$,

$$|\Phi_0\rangle = \mathrm{e}^{-\kappa} |0\rangle. \tag{1.28}$$

Here,

$$\kappa = \sum_{PQ} \kappa_{PQ} a_P^\dagger a_Q \tag{1.29}$$

is an anti-Hermitian operator (i.e. $\kappa_{PQ}^* = -\kappa_{QP}$) whose matrix elements κ_{PQ} are the orbital rotation coefficients and will be the variational parameters. The anti-Hermiticity of κ ensures that the orbital rotation $\mathrm{e}^{-\kappa}$ is unitary. Here we use the convention that capital P, Q denote spin-orbital indices.

With $\langle \Phi_0 | \Phi_0 \rangle = 1$, we want

$$E_{\mathrm{HF}} = \min_{\kappa_{PQ}} \langle \Phi_0 | \hat{H} | \Phi_0 \rangle. \tag{1.30}$$

This is a nonlinear equation, and hence the parameters κ_{PQ} must be determined iteratively. Since $\mathrm{e}^{-\kappa}$ is unitary, the orthonormality of the spin-orbitals is preserved, $\langle \phi_P | \phi_Q \rangle = \delta_{PQ}$.

The HF method can also be understood as a mean-field theory. That is, we treat the N-body problem as N one-body problems, where a single electron is in an effective

---

*More precisely, the HF approximation is to treat the wave function as a single configuration. In the context of this dissertation, that configuration will always be a SD.

"averaged" potential from the other $N - 1$ electrons. This effective one-body Hamiltonian is known as the Fock operator,

$$f = \sum_{PQ} f_Q^P a_P^\dagger a_Q. \tag{1.31}$$

The Fock operator f may be written as

$$f = h + V_{\text{eff}} \tag{1.32}$$

where the Fock potential's matrix elements (in spatial-orbital basis) are $V_{\text{eff},pq} = \sum_{i \in \text{occ}} (V_{ii}^{pq} - V_{iq}^{pi})$.

The Hartree-Fock orbitals, which ultimately dictate the parameters for the minimisation procedure, are determined by diagonalising the Fock matrix,

$$f_q^p = \epsilon_p \delta_q^p \tag{1.33}$$

where the eigenvalues ϵ_p are known as the orbital energies.

Since f itself depends on the orbitals, the solution must be determined self-consistently. Hence HF is often referred to as a self consistent field (SCF) method.

Often, the orbitals for the $\uparrow$ and $\downarrow$ electrons are restricted to be the same. This is known as the restricted Hartree-Fock (RHF) method. Otherwise, we might instead use the unrestricted Hartree-Fock (UHF) method, where they are treated independently. However, since this breaks spin symmetry, we might have spin contamination. Restricted open-shell Hartree-Fock (ROHF) is a variant that can treat open-shell systems which RHF otherwise would not be able to treat while remaining an eigenstate of the S^2 (total spin) operator.

1.5.1. Correlation Energy

The correlation energy (for a given basis set) is defined as the difference between the exact energy and the HF energy,

$$E_{\text{corr}} = E_{\text{exact}} - E_{\text{HF}}. \tag{1.34}$$

It should be noted, however, that despite this name the HF method captures Pauli exchange, and hence does actually correlate electrons. Nevertheless, we typically neglect this when we speak of "correlation energy".

Similarly, the Coulomb hole is defined as the remaining part of the wave function that is not captured by the HF method,

$$\Psi_{\text{hole}} = \Psi_{\text{exact}} - \Psi_{\text{HF}}. \tag{1.35}$$

1.5.2. The Roothaan-Hall Equations

The Roothaan-Hall equations are a matrix representation of the HF approximation.[30,31] Given a basis set, typically GTOs, the Roothaan-Hall equations are given by the generalised eigenvalue problem,

$$FC = SC\epsilon \tag{1.36}$$

where C is a matrix of coefficients, F is the Fock matrix (which depends on C), S is the overlap matrix (which reduces to the identity matrix for orthonormalised bases), and ϵ is a diagonal matrix of orbital energies. The Roothaan-Hall equations are obtained by expanding the unknown molecular orbitals (MOs) in a basis of known functions, typically atomic orbitals (AOs).

Since this representation is in matrix form, instead of in terms of derivatives and integrals, it is more amenable to conventional computational techniques.

1.6. Post-Hartree-Fock Methods

While the HF method is convenient and can be used efficiently to solve for many electrons, it is oftentimes not sufficient to describe the electronic structure of a molecule. To account for the remaining correlation, numerous post-Hartree-Fock methods (that is, methods run after HF) have been formulated, some of which are described in this section.

Correlation energy is typically categorised as either dynamical or static, although their effects are not mutually exclusive and so sometimes the distinction can be hazy.

Dynamical correlation is remaining correlation that arises due to the instantaneous repulsion from the motion (dynamics) of the electrons.

Static correlation, on the other hand, is related to the degeneracy or near-degeneracy of configurations. In general, to accurately describe a system with strong static correlation, a so-called multireference method is needed. As will be seen in chapter 5, static correlation plays a big role for molecules at dissociation.

We therefore have two "axes" on which to approach the exact solution to equation 1.5:

– The basis set chosen, where going to a larger basis set (in a systematic way) improves the result, e.g. cc-pVDZ to cc-pVTZ to cc-pVQZ, until eventually reaching the CBS limit.

– The "hierarchy of theories" where the more accurate the method, the closer to the full configuration interaction (FCI) limit (discussed below). Many post-Hartree-Fock methods can be systematically improved, for example by considering additional excitations in a coupled cluster method.
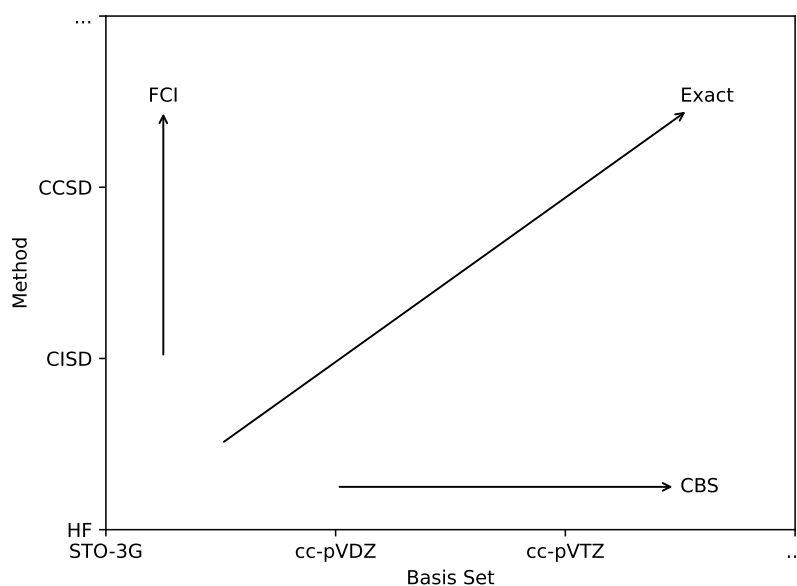
Figure 1.2. A sketch for the hierarchy of theories and basis sets. There are two axes on which to systematically improve until reaching the exact solution to the time-independent Schrödinger equation under the BOA: increasing basis set size (x axis) and increasing the level of theory (y axis). For most of this dissertation, we assume to be at or close to the highest level of theory (the FCI limit), as we mostly employ FCIQMC. We therefore focus on methods of rapidly converging to the CBS limit without the inherently expensive process of adding more basis functions.

1.6.1. Configuration Interaction

Configuration interaction (CI) is perhaps the most conceptually straightforward post-Hartree-Fock method. Instead of using a single SD, we approximate the wave function as a linear combination of SDs. That is,

$$|\Phi_{\mathrm{CI}}\rangle = \sum_p c_p |D_p\rangle \tag{1.37}$$

where $|D_p\rangle$ is a SD. This may alternatively be written[*]

$$|\Phi_{\mathrm{CI}}\rangle = (\mathbb{1} + \sum_\sigma \sum_{ia} c_i^a \hat{a}_{a\sigma}^\dagger \hat{a}_{i\sigma} + \sum_{\sigma\tau} \sum_{ijab} c_{ij}^{ab} a_{a\sigma}^\dagger a_{b\tau}^\dagger a_{i\tau} a_{j\sigma} + ...) |\Phi_0\rangle \tag{1.38}$$

where the number of terms in the sum depends on how many excitations we consider, i.e. how far we truncate the CI expansion. The energy and wave function are then found as the lowest eigenvalue and corresponding eigenvector, respectively, of the Hamiltonian written in the basis of these SDs. This can be done, for example, by exact diagonalisation. If all possible excitations are considered, then we have reached the CBS limit, and the solution is exact within that basis set.
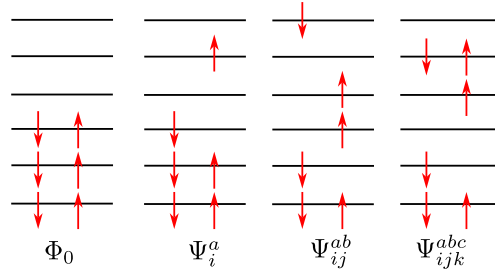

Figure 1.3. Illustration of a RHF reference and examples of single-, double-, and triple-excitations. The method CISDT, for example, would consider all determinants of the forms illustrated.

There are some important limitations to this method, however. Perhaps most notable is that the truncated CI method is not necessarily size consistent. For two noninteracting systems A and B, we expect the composite system $A + B$ to have energy equal to the sum of that for A and B, i.e. $E_{A+B} = E_A + E_B$. However, within a truncated CI method, this is not the case,[†] due to the additive ansatz 1.38.

Additionally, if we wish to go to the FCI limit and ensure size consistency and accurate results, we once again have bad scaling, as discussed in section 1.4. However, there exist sparse diagonalisation routines that need not store the entire matrix H, such as the Davidson,[32] Lanczos,[33] and Arnoldi[34] (for non-Hermitian H) methods, which target the

---

[*]We adopt the common notation convention where $i, j, k, ...$ denote occupied orbitals, $a, b, c, ...$ virtual (or unoccupied) orbitals, and $p, q, r, ...$ unspecified.

[†]Of course, in FCI we necessarily have size consistency, since we have the exact solution for the basis set.

ground state and need to store only a fraction of the matrix. Still, these scale with the size of the eigenvector, which is still prohibitively expensive for large systems. A more practical approach to reaching FCI accuracy will be discussed later in section 3.4, where we discuss the FCIQMC algorithm.

1.6.2. Multi-Configurational Self-Consistent Field

As a generalisation of the HF method, multiconfigurational self-consistent field (MCSCF) is particularly well-suited for problems with static correlation.[5,35,36] As in HF, we minimise the electronic energy with respect to variational parameters; however, now we also simultaneously minimise with respect to expansion coefficients for a set of configurations (e.g. SDs). i.e. consider the MCSCF wave function

$$|\kappa, \boldsymbol{c}\rangle = e^{-\kappa} \sum_i c_i |D_i\rangle, \tag{1.39}$$

where κ is familiar from HF, and the c_i coefficients familiar from CI. In MCSCF, we optimise

$$E_{\text{MCSCF}} = \frac{\langle \kappa, \boldsymbol{c}| H |\kappa, \boldsymbol{c}\rangle}{\langle \kappa, \boldsymbol{c}|\kappa, \boldsymbol{c}\rangle}. \tag{1.40}$$

Like in HF, the variational parameters appear nonlinearly and must be optimised iteratively. Determining an appropriate set $\{|D_i\rangle\}$ tends to be challenging, and even for small systems generating an MCSCF wave function can prove intractible.

Complete Active Space Self-Consistent Field

One "flavour" of MCSCF that has proven particularly successful is complete active space self-consistent field (CASSCF).[37–40] In CASSCF, instead of inspecting individual configurations, we consider a set of configurations that satisfy a set of criteria. In particular, we partition the orbitals into three sets: the core, active, and virtual (unoccupied) regions. The core orbitals, as alluded to in section 1.2.2, are approximated to be doubly occupied. The virtual orbitals are approximated to always be unoccupied. The active orbitals are the remaining orbitals, which can have occupations between 0 and 2.

The MCSCF expansion is found by considering all possible excitations of the active electrons in the active space. Notice that in the limit of an empty active space, we recover the HF method, and in the limit of an active space containing all orbitals, we recover the FCI method.

1.6.3. Perturbation Theory

Perturbation theory is a set of approximate mathematical methods for solving problems involving small disturbances (perturbations) to a problem with a known solution (the unperturbed problem). If these perturbations are not too large, then the solution of the

perturbed problem is close to that of the unperturbed problem, and can be expressed as the solution of the unperturbed problem plus some corrections. Perturbation theory fails, however, if the perturbation is large.

Rayleigh-Schrödinger

Perturbation theory applied to time-independent problems is sometimes referred to as Rayleigh-Schrödinger perturbation theory.[12,41,42] Consider a Hamiltonian,

$$H_{\mathrm{PT}} = H_0 + \lambda H', \tag{1.41}$$

where H_0 is referred to as the unperturbed Hamiltonian, λ is an arbitrary (real) parameter controlling the strength of the perturbation, and H' is the perturbation. We assume that we know the exact solution to H_0, such that we have all eigenstates $\{\left|\Psi_n^{(0)}\right\rangle\}$ and their corresponding eigenvalues $\{E_n^{(0)}\}$.

To obtain the solution to the true Hamiltonian H, we expand in terms of the perturbation λ,

$$\left|\Psi_n\right\rangle = \sum_{k=0}^{\infty} \lambda^k \left|\Psi_n^{(k)}\right\rangle \tag{1.42}$$

and

$$E_n = \sum_{k=0}^{\infty} \lambda^k E_n^{(k)}. \tag{1.43}$$

Since in this work we are generally interested in the ground state, we are typically targeting $n = 0$.

By inserting these expansions into the time-independent Schrödinger equation, we equate terms of the same order in λ, which leads to the corrections to the energy and wave function. We may then truncate the expansion to some order. This allows for a systematically improvable result.

Møller-Plesset

Møller-Plesset perturbation theory[43] is a special case of Rayleigh-Schrödinger perturbation theory, and is the variant most commonly seen in quantum chemistry.

In Møller-Plesset perturbation theory, the unperturbed Hamiltonian is chosen to be the Fock operator 1.31. The perturbed Hamiltonian is known as the fluctuation potential and is the difference between the true Coulomb interaction and the effective one-electron potential discussed in section 1.5.

That is,

$$H_0 = f, \quad H' = \sum_{i<j} r_{ij}^{-1} - V_{\mathrm{eff}} = H - H_0. \tag{1.44}$$

If we apply Rayleigh-Schrödinger perturbation theory, we find the zeroth order wave function to be the HF wave function,

$$\left| \Psi_0^{(0)} \right\rangle = \left| \Phi_0 \right\rangle, \quad f \left| \Phi_0 \right\rangle = \underbrace{\sum_i \epsilon_i}_{E_0^{(0)}} \left| \Phi_0 \right\rangle \tag{1.45}$$

where we have also identified the zeroth-order energy as the sum of orbital energies $E^{(0)} = \sum_i \epsilon_i$. With a bit more busywork, we get

$$E^{(1)} = \left\langle \Psi_0^{(0)} \right| H' \left| \Psi_0^{(0)} \right\rangle = \left\langle \Phi_0 \right| H' \left| \Phi_0 \right\rangle \tag{1.46}$$

$$E^{(2)} = \sum_{n>0} \frac{| \left\langle \Psi_n^{(0)} \right| H' \left| \Psi_0^{(0)} \right\rangle |^2}{E_0^{(0)} - E_n^{(0)}} = - \sum_{I>J, A>B} \frac{| \left\langle \Phi_{IJ}^{AB} \right| H' \left| \Phi_0 \right\rangle |^2}{\epsilon_A + \epsilon_B - \epsilon_I - \epsilon_J}, \tag{1.47}$$

where $\left| \Phi_{IJ}^{AB} \right\rangle = a_A^\dagger a_B^\dagger a_I a_J \left| \Phi_0 \right\rangle$ is a doubly-excited state with respect to the HF wave function, and $\epsilon_A + \epsilon_B - \epsilon_I - \epsilon_J = E_0^{(0)} - E_n^{(0)}$ is the energy difference between two eigenstates of the Fock operator.

Some features worth noting about Møller-Plesset perturbation theory:

- The sum of zeroth- and first-order energies is the HF energy: $E_{\mathrm{HF}} = E_0^{(0)} + E_0^{(1)}$.

- By the variational principle, $E_0^{(0)} < E_n^{(0)}$ for all $n > 0$ (except for degenerate ground states), so $E_0^{(2)} < 0$, i.e. the energy always decreases.

- If the ground state is degenerate, the term diverges. It might be possible to lift the degeneracy by a change of basis, however.

- If the HF solution is already a good approximation for the system, then Møller-Plesset perturbation theory can provide surprisingly good (and size consistent) results. Hence, this method is typically only applicable to single-reference problems (i.e. those without strong static correlation).

While higher-order approaches exist and see use, the most popular is second-order Møller-Plesset perturbation theory (MP2), due to its excellent compromise between cost and accuracy. That is, for only a little extra work after a successful HF, applying MP2 can improve our results considerably.

1.6.4. Coupled Cluster Theory

The lack of size consistency in CI theory arises from the linear ansatz in section 1.6.1. By instead using an exponential ansatz, we arrive at one of the most successful theories in electronic theory, coupled cluster (CC).[44–47] While the methods presented here are

inherently single-reference and build on a single SD, multireference generalisations to CC theory exist and are the study of active research.[48–53] Therefore, the presented "flavours" of CC are generally expected to fail for systems with strong static correlation.

Standard Coupled Cluster

To ensure a multiplicatively-separable wave function, we use a multiplicative ansatz,

$$|\Psi_{\mathrm{CC}}\rangle = \mathrm{e}^T |\Phi_0\rangle, \tag{1.48}$$

where $|\Phi_0\rangle$ is the reference (typically HF) wave function, and T is the cluster operator. The cluster operator is defined $T := T_1 + T_2 + T_3 + \cdots$, where T_n is the n-body cluster operator, made up of all possible n-body excitations, for example, $T_1 = \sum_\sigma \sum_{i,a} t_{ia} a_{a\sigma}^\dagger a_{i\sigma}$ for the one-body cluster operator, and $T_2 = \sum_{\sigma\tau} \sum_{i<j,a<b} t_{ij}^{ab} a_{a\sigma}^\dagger a_{b\tau}^\dagger a_{j\tau} a_{i\sigma}$ for the two-body cluster operator. The values $t_{i\cdots}^{a\cdots}$ are referred to as the cluster amplitudes.

It is worth noting that CC and CI are identical, differing only in their parametrisation, when neither are truncated. They both provide FCI-level accuracy.

By inserting equation 1.48 into the time-independent Schrödinger equation and pre-multiplying by e^{-T}, we obtain

$$\langle\Phi_0| \mathrm{e}^{-T} H \mathrm{e}^T |\Phi_0\rangle = E \tag{1.49}$$

$$\left\langle\Phi_{ij\cdots}^{ab\cdots}\right| \mathrm{e}^{-T} H \mathrm{e}^T |\Phi_0\rangle = 0 \tag{1.50}$$

where $\left|\Phi_{ij\cdots}^{ab\cdots}\right\rangle$ is a n-body excitation with respect to the reference wave function.

Note, however, that these equations do not rely on the Variational Principle. Instead of minimising a functional, we solve equations 1.50 for the cluster amplitudes, which give the energy. While we still cannot reasonably always include all excitations in the cluster operator T, due to the multiplicative ansatz, even if we truncate the cluster operator, we get size-consistent results. In particular, any truncated CC wave function ansatz will contain contributions from all determinants in Fock space. The simplest CC truncation commonly used is the coupled cluster with single and double excitations (CCSD) method. In a sense, CCSD and any other truncated CC method is an approximation for FCI where the coefficients are not approximated to be zero (like in CI) but instead the higher-order excitations terms are generated from the lower-order excitation terms in a size-consistent fashion.

CC theory owes much of its success to its high accuracy at a reasonable cost. CCSD and coupled cluster with single, double and triple excitations (CCSDT) in particular scale as $\mathcal{O}(M^6)$ and $\mathcal{O}(M^8)$ respectively, with M the number of spatial orbitals. Every additional set of excitations incurs an additional factor of $\mathcal{O}(M^2)$ scaling.

Distinguishable Cluster Theory

The more recently-developed distinguishable cluster with single and double excitations (DCSD) method's central philosophy is removing irrelevant or nonphysical terms while maintaining many of CC's desireable properties, such as size consistency.[54]

This method was shown to give quantitatively better results than standard CC for a variety of systems, and qualitatively correct results for systems where CC can fail spectacularly, such as the N_2 binding curve. Since its first publication, distinguishable cluster (DC) has been extended to explicitly-correlated F12 theory,[55] equation of motion (EOM),[56] tailored coupled cluster,[57] and more.

Combining Coupled Cluster with Perturbation Theory

It is possible to avoid the computationally-expensive scaling of CC somewhat, by treating some terms perturbatively. The most famous of these is the coupled cluster with single and double excitations and perturbative triples (CCSD(T)) method,[58] in which the single- and double-excitations are treated fully, whereas the triple-excitations are treated as perturbations, as in section 1.6.3.

CCSD(T) scales as $\mathcal{O}(M^7)$ and for the price provides extremely accurate results. The perturbative correction is a step that is applied only once at the end of a CCSD calculation. This is why sometimes this method is referred to as the "gold standard" of quantum chemistry.

Explicitly Correlated Methods

With the advent of modern computers combined with a vast array of sophisticated algorithms from which to choose, *ab initio* quantum chemistry has become a tremendously powerful tool, going beyond the study of small atoms, to molecules and solids, and are among the most effective and systematically improvable techniques to date. Nevertheless, convergence to the CBS limit is notoriously slow.

In particular, consider the popular basis set family developed by Dunning and coworkers, correlation-consistent polarised valence with X-zeta quality (X=D, T, Q, 5, 6, ...) (cc-pVXZ).[14,21,59–61] The size of these basis sets scale as $M \in \mathcal{O}(X^3)$, and since for standard post-Hartree-Fock discussed in chapter 1.6 we require four-index integrals, our computation time will scale at best as $t \in \mathcal{O}(X^{12})$.[62]

Meanwhile, the CBS correlation error scales as $\epsilon \in \mathcal{O}(X^{-3})$[16,18] or $\epsilon \in \mathcal{O}(M^{-1})$,[63] resulting in time $t \in \mathcal{O}(M^{-\frac{1}{4}})$. Thus, the methods discussed so far come with the painful cost of requiring very large basis sets to approach high-accuracy results.

Explicitly correlated methods are a class of electronic structure methods specifically designed to address this unfavourable scaling by explicitly including the interelectronic distance r_{12}, and is the subject of this chapter. As the R12/F12 family of methods is the most mature of the explicitly correlated electronic structure methods, many reviews focusing on this topic already exist. This chapter in particular is in large part based on references [62, 64–66].

2.1. The Cusp Conditions

Consider two charged point particles in a system described by the Hamiltonian of equation (1.5). By the Schrödinger equation, the local energy

$$E_L := \frac{H\Psi}{\Psi} \tag{2.1}$$

must be constant in the exact solution. However, when these two particles coalesce, i.e. $r_{12} \to 0$, the Coulomb potential, r_{12}^{-1}, diverges. Thus, for the local energy to be constant, we must have that near coalescence points, the kinetic energy exactly cancels the Coulomb energy. A more formal treatment of this argument leads to the (antiparallel)

electron-electron Kato cusp condition,[67]

$$\left.\frac{\widetilde{\partial\Psi}}{\partial r_{12}}\right|_{r_{12}\to0} = \frac{1}{2}\Psi(r_{12}=0) \tag{2.2}$$

where the tilde represents spherical averaging.

This cusp condition was also later generalised.[68,69] Early literature on the subject suggested that the success of explicitly correlated methods were due to the superior description of short-range correlation effects, and in particular in their much more faithful capturing of the cusp conditions like equation 2.1.[70–73] However, further study found that the correlation error from a bad description of the wave function around a small sphere centred on the cusp is actually negligible.[62,74–76] Instead, the success of explicitly correlated methods is actually due to the superior description of the overall shape and size of the Coulomb hole, which has a radius on the order of the atomic radius.

To understand why gaussian-type basis sets fail so spectacularly at capturing the cusp behaviour, it is instructive to consider a simpler example, like that of approximating $|x|$ by its Fourier decomposition. Such an illustration is found in figure 2.1. It is also worth noting that STOs do not suffer as badly from this limitation,[77] but as discussed in chapter 1.4.1, they are unsuccessful due to their lack of practicality.*

2.2. Hylleraas Methods

Almost 30 years prior to Kato's landmark paper describing cusps in the analytical form for the wave function, there was already work being done to understand the significance of including r_{12} terms in the wave function. Instead of being motivated by rigorous mathematics like Kato's, Slater was motivated by studies in the He atom. In particular, he tried to construct a wave function which faithfully represents both the core region as well as the Rydberg limit (i.e. a highly excited atom where the electron is very far from the nucleus).[77–80] This led him to suggest multiplying the wave function by a factor

$$e^{-2(r_1+r_2)+r_{12}/2}, \tag{2.3}$$

which can easily be shown to satisfy the cusp equation (2.2).

However, the first successful explicitly correlated electronic structure calculation is typically attributed to Hylleraas,[65] where he aimed to improve convergence of orbital expansions for helium.[81,82] In this method, the coordinates $s := r_1 + r_2$, $t := r_1 - r_2$ and

---

*Note that STOs do not suffer from the electron-nuclear cusp, which is in any case not the focus of many explicitly correlated methods, such as F12, as discussed later in this chapter. Even with STOs, we have electron-electron cusps.
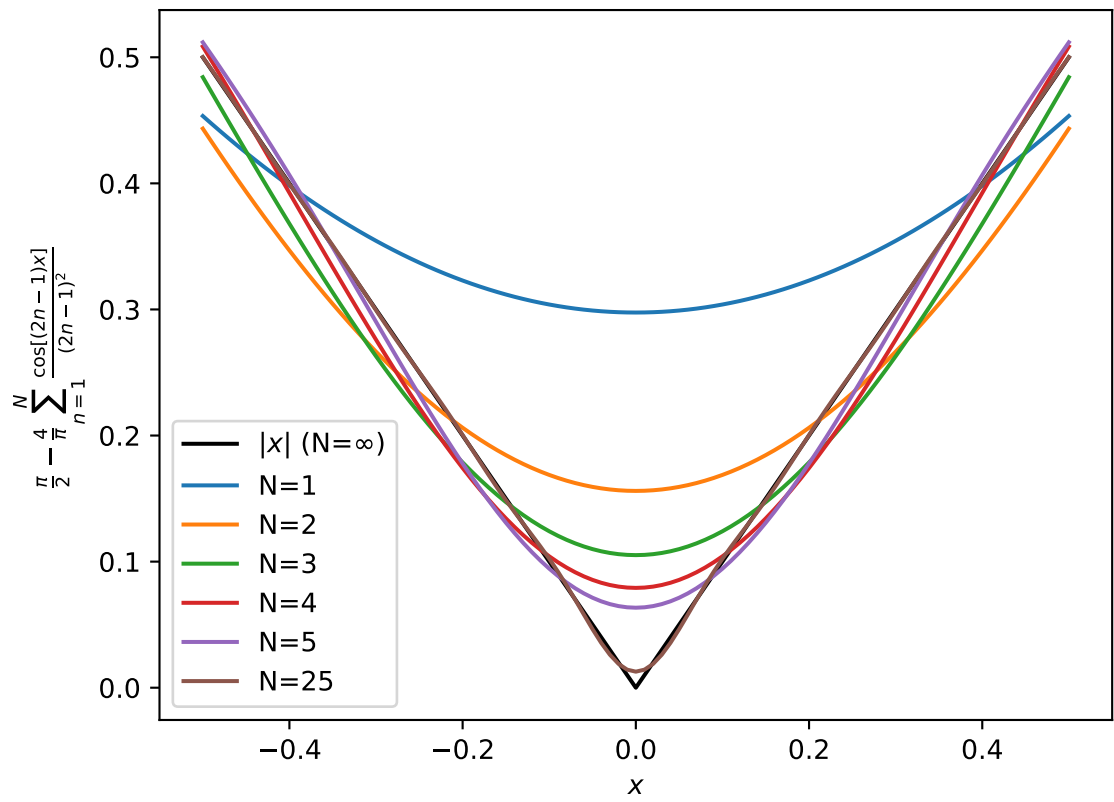
Figure 2.1. A toy example of the Coulomb cusp. Here, the Fourier expansion $x \approx \frac{\pi}{2} - \frac{4}{\pi} \sum_{n=1}^{N} \frac{\cos[(2n-1)x]}{(2n-1)^2}$ is plotted for a few values of N, including the exact solution. As can be seen, even for many terms, the Fourier expansion is a poor descripter in the cusp region. Indeed, the only way to describe it exactly is with an infinite number of terms.

$u := r_{12}$ are used to construct the wave function

$$\Psi_N(s, t, u) = e^{-\alpha s} \sum_{k}^{N} c_k s^{l_k} t^{2m_k} u^{n_k}. \tag{2.4}$$

In particular, using only three terms ($N = 3$), and variationally optimising for the parameters c_1, c_2, c_3, Hylleraas was able to reach within 1.3 millihartree from the exact result.

Since then, there was rapid development on this approach and combining it with CI (which came to be known as the CI-Hyl methods).[83–91] In CI-Hyl methods, the wave function is expanded as in CI,

$$\Psi = \sum_{k} c_k \Phi_k \tag{2.5}$$

where

$$\Phi_k = \mathcal{A} r_{ij}^{\nu_k} \prod_{i} \chi_{k_i}(\boldsymbol{x}_i) \tag{2.6}$$

where χ_k is a spin-orbital basis and $\mathcal{A}$ is the antisymmetriser operator.

However, CI-Hyl methods were to eventually fall out of favour. This is because the expansions involve exceedingly difficult integrals involving many electrons and over products of correlation factors. This significantly restricts the tractibility and scalability of the method, and it has since largely gone unused.

2.3. Explicitly Correlated Gaussians

Boys[92] and Singer[93] independently introduced gaussian basis functions with explicit correlation for calculations on molecules.[94] These methods are referred to as Explicitly Correlated Gaussians (ECGs), or in the case of functions of two electrons, Gaussian-type Geminals (GTGs). A spherical GTG may be written (with nuclar coordinates $\boldsymbol{R}_1$ and $\boldsymbol{R}_2$) as

$$g(\boldsymbol{r}_1, \boldsymbol{r}_2) = \exp\left(-\zeta_1|\boldsymbol{r}_1 - \boldsymbol{R}_1|^2 - \zeta_2|\boldsymbol{r}_2 - \boldsymbol{R}_2|^2 - \gamma|\boldsymbol{r}_1 - \boldsymbol{r}_2|^2\right). \tag{2.7}$$

This can be interpreted as two s-type orbitals and an interelectronic correlation factor $e^{-\gamma r_{12}^2}$. While they do not have the correct cusp behaviour, a linear combination of GTGs approximately capture the electron-electron cusp. This is similar to how gaussian basis functions do not individually capture the electron-nuclear cusp, but a linear combination of them approximately does.

One major strength of this approach is that all integrals have closed-form algebraic expressions,[95] and avoids nonlinear optimisation.[96,97] ECG methods have been extended to post-Hartre-Fock methods, such as MP2,[98,99] and methods to avoid its many difficult integrals have also been developed.[100–104]

While not as popular as F12 methods (see section 2.4), ECGs have been used for highly accurate variational calculations,[105] as well as for applications outside of standard electronic structure theory, such as bosons,[106] positronium (a bound state of an electron and a positron),[107] and non-Born-Oppenheimer systems.[108]

2.4. F12 Methods

The most influential class of explicitly correlated methods to date are the F12 methods. The core principle of these methods is to augment the wave function from a conventional (typically SD) basis with an explicitly-correlated correction, called the F12 (or R12) correction. The original formulation[109] parametrised a two-electron system (such as helium) wave function as

$$|\Psi\rangle = (1 + tQ_{12}F_{12}(r_{12}))\,|\Psi_0\rangle + \sum_p t_p\,|\Psi_p\rangle \qquad (2.8)$$

where Ψ_0 is a reference wave function (such as HF), $|\Psi_p\rangle$ are excited-state SDs, t, t_p are parameters ("amplitudes") to be optimised, and

$$Q_{12} := \sum_{\alpha\beta} |\alpha\beta\rangle \langle\alpha\beta| \qquad (2.9)$$

is referred to as a "strong orthogonality projector", which ensures the $F_{12}(r_{12})$ term is orthogonal to the reference and singly-excited determinants. Here α, β refer to virtual orbitals in the formally complete basis. The fact that, due to Q_{12}, the $F_{12}(r_{12})$ term commutes with the standard excitation operators aids in including this explicit correlation into conventional electron correlation methods. This geminal term is added as a correction to the standard wave function.

In the original formulation, $F_{12}(r_{12}) = r_{12}$ (hence the name "R12" or "F12" for the more general methodology). However, another popular choice is a Slater-type geminal[110] fitted to a linear combination of gaussians,[111]

$$F_{12}(r_{12}) = -\gamma^{-1}e^{-\gamma r_{12}} = -\gamma^{-1} + r_{12} - \frac{1}{2}\gamma r_{12}^2 + \cdots \approx \sum_i c_i e^{-\alpha_i r_{12}^2}. \qquad (2.10)$$

The choice of the length scale γ can either be optimised or (more typically) kept fix, although formally it is orbital dependent and a poor choice may result in a loss of accuracy.[112] Other choices for the correlation factor also exist.[113]

2.4.1. MP2-F12

Following the discussion in reference 66, and adopting their notation, we consider the MP2-F12 method as an illustrative example.

An alternative derivation of the MP2 equations from section 1.6.3 is to minimise the Hylleraas functional[114,115]

$$E^{(2)} = \min_{\Psi} \left\langle \Psi^{(1)} \middle| (H_0 - E^{(0)}) \middle| \Psi^{(1)} \right\rangle + 2 \left\langle \Psi^{(1)} \middle| H \middle| \Psi^{(0)} \right\rangle \tag{2.11}$$

where H_0 is the unperturbed Hamiltonians, and the (i) superscript denotes the order of the correction, as introduced in section 1.6.3.

The earliest generalisation of equation 2.8 was an ansatz for MP2,[116–120]

$$\left| \Psi^{(1)} \right\rangle = \frac{1}{2} \sum_{ij} \left(\sum_{ab} t_{ij}^{ab} \left| \Psi_{ij}^{ab} \right\rangle + \sum_{kl} t_{ij}^{kl} \sum_{\alpha\beta} \left\langle \alpha\beta \middle| Q_{12} F_{12}(r_{12}) \middle| kl \right\rangle \left| \Psi_{ij}^{\alpha\beta} \right\rangle \right), \tag{2.12}$$

$$Q_{12} = (1 - o_1)(1 - o_2) - v_1 v_2 \tag{2.13}$$

where we define the one-electron projection operators

$$o_m = \sum_i |\phi_i(\boldsymbol{r}_m)\rangle \langle\phi_i(\boldsymbol{r}_m)| \tag{2.14}$$

and

$$v_m = \sum_a |\phi_a(\boldsymbol{r}_m)\rangle \langle\phi_a(\boldsymbol{r}_m)| \tag{2.15}$$

such that o_m and v_m project onto the occupied and virtual orbitals, respectively, with

$$\langle\phi_i(\boldsymbol{r}_m)| \Omega |jk\rangle = \int \mathrm{d}^3 r_m \ \phi_i^*(\boldsymbol{r}_m)\Omega\phi_j(\boldsymbol{r}_1)\phi_k(\boldsymbol{r}_2) \tag{2.16}$$

for operator Ω. The first term in brackets in equation 2.12 are from the conventional MP2 wave function correction, and the extra term is a contraction of a formally-infinite set of double excitations, representing the F12 correction (again, orthogonal to the conventional wave function thanks to Q_{12}).

It was also shown that solving for the amplitudes t_{ij}^{kl} can be avoided,[21,111,121–123] and can instead be derived from the cusp conditions to give

$$t_{ij}^{kl} = \frac{3}{8}\delta_{ik}\delta_{jl} + \frac{1}{8}\delta_{il}\delta_{jk}. \tag{2.17}$$

This also tends to be more accurate, as it avoids the geminal basis set superposition error.[122,124]

2.4.2. Many-Electron Integrals

The Hylleraas functional, equation 2.11, will contain terms such as

$$\sum_{\alpha\beta} \langle\Phi_0| H \left|\Psi_{ij}^{\alpha\beta}\right\rangle \langle\alpha\beta| Q_{12}F_{12}(r_{12}) |kl\rangle$$

$$=2\langle ij| r_{12}^{-1}Q_{12}F_{12}(r_{12}) |kl\rangle - \langle jl| r_{12}^{-1}Q_{12}F_{12}(r_{12}) |kl\rangle. \qquad (2.18)$$

Plugging in the form for Q_{12} from equation 2.13, we get integrals of the form

$$\langle ij| r_{12}^{-1}o_1 F_{12}(r_{12}) |kl\rangle. \qquad (2.19)$$

For $F_{12}(r_{12}) = r_{12}$, this is equal to

$$\sum_{o} \langle ijo| r_{12}^{-1}r_{23} |olk\rangle, \qquad (2.20)$$

which is a three-electron integral. Here, o is an occupied orbital. Clearly, with more operators multiplied together in the kernel of these integrals, we might expect even higher-order integrals. This may seem distrastrous, as this would be a massive bottleneck.

However, it was found that by insertion of the resolution of identity (RI),[116,117,125] these integrals can be reduced to a linear combination of two-electron integrals. In particular, the strong projection operator becomes

$$Q_{12} =1 - \sum_{o}\sum_{\alpha}(|o\alpha\rangle\langle o\alpha| + |\alpha o\rangle\langle\alpha o|) + \sum_{o,o'}|oo'\rangle\langle oo'| - \sum_{a,b}|ab\rangle\langle ab|, \qquad (2.21)$$

and the discussed three-body integral becomes

$$\sum_{o} \langle ijo| r_{12}^{-1}r_{23} |olk\rangle =\delta_{ik}\delta_{kl}$$

$$- \sum_{o}\sum_{\alpha}(\langle ij| r_{12}^{-1} |o\alpha\rangle\langle o\alpha| r_{12} |kl\rangle + \langle ij| r_{12}^{-1} |\alpha o\rangle\langle\alpha o| r_{12} |kl\rangle)$$

$$+ \sum_{o,o'} \langle ij| r_{12}^{-1} |oo'\rangle\langle oo'| r_{12} |kl\rangle - \sum_{a,b} \langle ij| r_{12}^{-1} |ab\rangle\langle ab| r_{12} |kl\rangle. $$

$$(2.22)$$

In the RI approximation, the summation over the additional orbitals α is in principle infinite. In practice, of course, these must be finite. Initially, the basis in the RI expansion was set equal to the orbital basis (dubbed the "standard approximation").[116–118,125–130] However, normally larger bases are used, such as in the complementary auxiliary basis set (CABS) approach, where the RI basis is equal to the orbital basis, plus some "auxiliary" basis functions that are orthogonal to the orbital basis.[123,129] Many more important advances were since made, such as more efficient RI methods and ways of generating intermediate values have also been developed.[110,111,121,123,131–147]

Moreover, it has been found that the original $F_{12}(r_{12}) = r_{12}$ form requires larger auxiliary basis functions, since its long-range behaviour is unphysical.[110,121] Instead, functions such as that in equation 2.10 is used.

2.4.3. Higher-Order F12 Theories

F12 theory has been extended to many other theories besides MP2, such as CC, CASSCF, multireference configuration interaction (MRCI), complete active space perturbation theory to second order (CASPT2), density matrix renormalization group (DMRG), and FCIQMC.[66,138,148–163] The key change compared to conventional methods such as those in chapter 1 is in the excitation operators, particularly the double excitations E_{ij}^{ab}. In particular, for a multireference method with reference determinants $|\Phi_I\rangle$, the wave function may be parametrised as

$$|\Psi\rangle = |\Psi_{\text{conv}}\rangle + \frac{1}{2} \sum_I \sum_{ijkl} t_{klI}^{ij} \sum_{\alpha\beta} \langle\alpha\beta| Q_{12}^I F_{12}(r_{12}) |kl\rangle E_{ij}^{\alpha\beta} |\Phi_I\rangle, \qquad (2.23)$$

where $|\Psi_{\text{conv}}\rangle$ is the conventional wave function. In this way, we extend the excitations to include geminal terms in the (formally complete) auxiliary basis.

These F12 methodologies have been particularly successful for large systems, proving only marginally more expensive than the conventional methods in some cases.[124,164]

Since F12 methods work directly on the wave function, it is also natural to extend them to excited states, particularly for multireference methods.[152,161,162,165,166] However, since the geminal terms include only occupied orbitals, there is an inherent bias to the ground state.[152,153] Several means of ameliorating this have been proposed, such as combining F12 with response theory,[153] or by extending the geminal basis to include virtual orbitals.[154]

2.5. The Transcorrelated Method

One of the main focuses of this dissertation is TC. Hirschfelder was the first to propose a similarity-transformed Hamiltonian method in the 1960s,[167] which was further developed by Jankowski,[168,169] and later by Boys and Handy.[170–173]

The core concept is to similarity-transform the Hamiltonian using some invertible operator P, such that

$$\hat{H}_{\text{TC}} = P^{-1}\hat{H}P. \qquad (2.24)$$

For any eigenvalue and eigenfunction pair (E_i, Ψ) of $\hat{H}$, the transformed Hamiltonian $\hat{H}_{\text{TC}}$ has eigenfunction $\Phi := P\Psi$ with the same eigenvalue E_i. This is called isospectrality, and it allows us to solve $\hat{H}_{\text{TC}}$ to get the spectrum of $\hat{H}$.

However, note that there is no requirement that $\hat{H}_{\text{TC}}$ is Hermitian, and in fact it typically is not. Recall from section 1.3 that the variational principle relies on the Hermiticity of the operator. It is therefore nontrivial to make TC variational, and not all

conventional methods can be used to treat $\hat{H}_{\text{TC}}$ (although, as discussed in section 3.4, projector methods such as FCIQMC and CC may be used).

2.5.1. The Method of Boys and Handy

One of the earliest practical approaches, on which most modern variations build, is that of Boys and Handy.[170–173] In this methodology, we start with a Jastrow ansatz[174]

$$\Psi = \mathrm{e}^J \Phi. \tag{2.25}$$

In this way, TC shares similarities with variational Monte Carlo (VMC), to be discussed in section 3.2. Inserting 2.25 into the Schrödinger equation, we get

$$\hat{H}\mathrm{e}^J \Phi = E\mathrm{e}^J \Phi \implies \underbrace{\mathrm{e}^{-J}\hat{H}\mathrm{e}^J}_{\hat{H}_{\text{TC}}} \Phi = E\Phi, \tag{2.26}$$

i.e. similarity-transforming $\hat{H}$ with the operator $P \equiv \mathrm{e}^J$.

In the original formulation, Φ was chosen to be a single SD (HF), and J was chosen such that

$$J = \sum_{i<j} u(\boldsymbol{r}_i, \boldsymbol{r}_j) \tag{2.27}$$

where $u(\boldsymbol{r}_i, \boldsymbol{r}_j) = u(\boldsymbol{r}_j, \boldsymbol{r}_i)$ is a symmetric two-electron correlation function.

The Baker-Campbell-Hausdorff (BCH) expansion of $\hat{H}_{\text{TC}}$ truncates exactly to second order, i.e.

$$\hat{H}_{\text{TC}} := \mathrm{e}^{-J}\hat{H}\mathrm{e}^J = \hat{H} + [\hat{H}, J] + \frac{1}{2}[[\hat{H}, J], J]. \tag{2.28}$$

This may be rewritten

$$\hat{H}_{\text{TC}} = \hat{H} - \sum_{i<j} \hat{K}(\boldsymbol{r}_i, \boldsymbol{r}_j) - \sum_{i<j<k} \hat{L}(\boldsymbol{r}_i, \boldsymbol{r}_j, \boldsymbol{r}_k), \tag{2.29}$$

where, using the shorthand $u_{ij} := u(\boldsymbol{r}_i, \boldsymbol{r}_j)$,

$$\hat{K}(\boldsymbol{r}_i, \boldsymbol{r}_j) = \frac{1}{2}\left(\nabla_i^2 u_{ij} + \nabla_j^2 u_{ij} + \left(\nabla_i u_{ij}\right)^2 + \left(\nabla_j u_{ij}\right)^2\right) + \left(\nabla_i u_{ij} \cdot \nabla_i\right) + \left(\nabla_j u_{ij} \cdot \nabla_j\right), \tag{2.30}$$

and

$$\hat{L}(\boldsymbol{r}_i, \boldsymbol{r}_j, \boldsymbol{r}_k) = \nabla_i u_{ij} \cdot \nabla_i u_{ik} + \nabla_j u_{ji} \cdot \nabla_j u_{jk} + \nabla_k u_{ki} \cdot \nabla_k u_{kj}. \tag{2.31}$$

In second quantisation, the TC Hamiltonian can then be written

$$\hat{H}_{\text{TC}} = \sum_{pq} \sum_{\sigma} h_q^p a_{p\sigma}^\dagger a_{q\sigma} + \frac{1}{2} \sum_{pqrs} \left(V_{rs}^{pq} - K_{rs}^{pq} \right) \sum_{\sigma\tau} a_{p\sigma}^\dagger a_{q\tau}^\dagger a_{s\tau} a_{r\sigma}$$
$$- \frac{1}{6} \sum_{pqrstu} L_{stu}^{pqr} \sum_{\sigma\tau\lambda} a_{p\sigma}^\dagger a_{q\tau}^\dagger a_{r\lambda}^\dagger a_{u\lambda} a_{t\tau} a_{s\sigma},$$
(2.32)

where

$$h_q^p = \langle p| \, h \, |q\rangle \,,$$
$$V_{rs}^{pq} = \langle pq| \, r_{12}^{-1} \, |rs\rangle \,,$$
$$K_{rs}^{pq} = \langle pq| \, \hat{K} \, |rs\rangle \,,$$
$$L_{stu}^{pqr} = \langle pqr| \, \hat{L} \, |stu\rangle \,.$$
(2.33)

Particularly noteworthy are the three-electron integrals from $\hat{L}$ and the non-Hermitian terms from $\hat{K}$.

Boys and Handy's original choice for the form of u was

$$u(\boldsymbol{r}_i, \boldsymbol{r}_j) = \sum_{l,m,n}^{l+m+n \leq S} c_{lmn} \bar{r}_{ij}^l (\bar{r}_{Ii}^m \bar{r}_{Jj}^n + \bar{r}_{Ij}^m \bar{r}_{Ji}^n)$$
(2.34)

where S is some integer ($S = 6$ in reference 170)[175] and

$$\bar{r} = \frac{r}{1+r}.$$
(2.35)

In this form, electron-electron (e-e), electron-nucleus (e-n), electron-electron-nucleus (e-e-n) and electron-electron-nucleus-nucleus (e-e-n-n) terms appear when $m = n = 0$ and $l > 0$, $m, n > 0$ and $l = 0$, one of $m, n > 0$ and $l > 0$, or $m, n, l > 0$ respectively.

In the original formulation,[170] the following equations are approximately solved via quadrature for the orbital coefficients a_i and the coefficients c_{lmn} in u.

$$\langle \Phi | \, (\hat{H}_{\text{TC}} - E) \, |\Phi\rangle = 0$$
(2.36)

$$\langle \partial \Phi / \partial a_i | \, (\hat{H}_{\text{TC}} - E) \, |\Phi\rangle = 0$$
(2.37)

$$\langle \Phi | \, (\hat{H}_{\text{TC}} - \hat{H}_{\text{TC}}^\dagger) \, |\Phi\rangle = 0,$$
(2.38)

and the energy is calculated via projection,

$$\langle \Phi | \, \hat{H}_{\text{TC}} \, |\Phi\rangle \,.$$
(2.39)

Equation 2.38 in particular is to make the TC Hamiltonian as close to Hermitian as possible for orbital optimisation.

This formulation was further refined by Handy.[176] He notice that since $\hat{H}_{\mathrm{TC}}$ is non-Hermitian, we cannot directly minimise the energy. Instead, he minimised the transcorrelated variance,

$$\sigma^2 = \langle [(\hat{H}_{\mathrm{TC}} - E)\Phi]^2 \rangle. \tag{2.40}$$

Furthermore, equation 2.34 is not the only possible choice for the Jastrow factor, and in particular Handy chose a form with simpler analytical integrals.

2.5.2. Modern Resurgence

For several decades TC received little attention, though some more developments were made, such as calculations using a fixed Jastrow factor,[177] more efficient integration,[178,179] multireference Φ,[180] among others.[181–183] More recently, there has been a renewal of interest in TC,[175,181–223] notably the demonstrations that the TC Hamiltonian may be treated by various electronic structure methods, to be discussed more in section 3.4.7.[175,185–188,196,203–205,213,214,217] TC has since been applied to a variety of systems, and in the context of quantum computing.[193,201,207,208,215] After an overview of stochastic methods, the rest of this dissertation will be focused on discussing recent developments in TC.

In typical modern TC workflows, Φ is initially kept fixed (e.g. to the HF determinant, though it need not be this), and then the parameters of the Jastrow factor, J, are optimised. The integrals for the similarity-transformed Hamiltonian $\hat{H}_{\mathrm{TC}}$ are then calculated, and finally an electronic structure method such as FCIQMC is applied to $\hat{H}_{\mathrm{TC}}$, solving for E and in effect "recalculating" Φ, which was previously kept fixed. The choice of J has ranged from especially simple[188,194,216] to much more sophisticated.[175,224,225]

TC is still behind F12 in terms of its performance to cost ratio, but F12 is ahead several decades of development. However, TC shows promise thanks in part to the high degree of flexibility owing to the Jastrow factor. Moreover, the additive ansatz of F12 leads to many-electron integrals, whereas the multiplicative ansatz of TC leads to at most three-electron integrals.* This is ameliorated in F12 via the use of RI approximations and in TC via other approximations, like discussed in section 4.9.1.

---

*This assumes the Jastrow factor may be written $J = \sum_i u_i(\boldsymbol{r}_i, \boldsymbol{r}_j)$.

Monte Carlo Methods

Monte Carlo (MC) methods are a class of numerical methods that use random sampling to numerically solve problems. It has found applications in an impressive range of fields, from physics to finance.[226–231] It is particularly useful for problems with high dimensionality, where deterministic methods are often impractical. In quantum chemistry and physics, since a 'dimension' can refer to any degree of freedom, high-dimensional problems are commonplace, and so MC methods are a natural choice.

While the name *Monte Carlo* was coined by Stanislav Ulam, after the famous casino in Monaco,[232,233] the foundational concept was already developed in the 18th century by the French mathematician Georges-Louis Leclerc, Comte de Buffon. As one of the earliest example applications, in the Buffon needle problem, one can randomly toss needles onto a lined sheet of paper and determine π.[234–236]

Monte Carlo methods is a broad term, and as such it is not possible to give a comprehensive overview in a single chapter, and there exist many reviews and textbooks on MC and related topics.[237–242] Here, we will focus on only a few concepts particularly relevant for this dissertation, largely following reference 230 and the relevant chapters of reference 231.

3.1. Classical Monte Carlo Methods

We start our discussion with classical MC methods. While there are numerous possible applications, notably in molecular dynamics,[237] here we restrict ourselves to the topic of Monte Carlo integration. In particular, we consider the classical textbook problem of calculating the value of π, then we provide a more rigorous framework.

3.1.1. A Very Bad Game of Darts

If we imagine throwing darts at a dartboard randomly, we can approximate π. If the radius of the circle is r, then its area is πr^2. The area of the square circumscribing the circle is $4r^2$. Therefore, the ratio of the area of the circle to the area of the square is $\pi/4$. If we randomly sample a point in the square ("throw a dart"), the probability that the point is inside the circle is proportional to its area. Since we sample inside the square,

the probability of landing inside the circle is

$$P(\text{inside circle}) = \frac{\pi r^2}{4r^2} = \frac{\pi}{4}. \tag{3.1}$$

Therefore, if we sample a large number of points, the ratio of the number of darts that land inside the unit circle to the total number of darts, we can approximate the probability distribution P and thus get an estimate for π as

$$\pi \approx 4\frac{N_{\text{in}}}{N_{\text{out}}}. \tag{3.2}$$
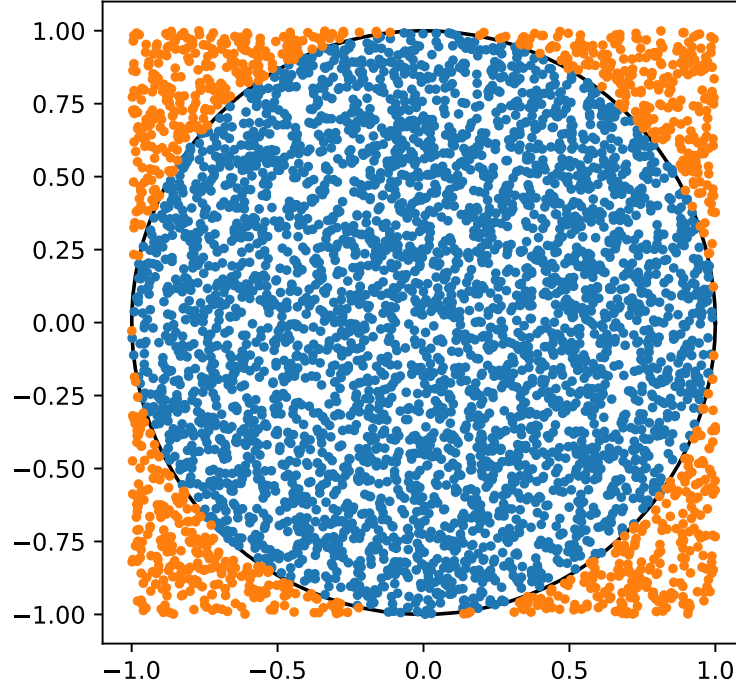
This is illustrated in figure 3.1, and captures the core philosophy of MC methods.

Figure 3.1. Our "game of darts". Points inside the unit circle are coloured blue whereas points outside are orange. Using stochastic sampling, this naive approach uses 5000 randomly generated coordinates $(x, y) \in [-1, 1] \times [-1, 1]$ to approximate $\pi \approx 4N_{\text{in}}/N_{\text{out}} \approx 3.1464$.

Controlling the stochastic error is critical in MC methods. To estimate the reliability of our estimate, we can determine the standard deviation of the estimate. Our sampling scheme is a binomial distribution, with $p = \pi/4$. Hence,

$$\sigma_\pi^2 = \text{Var}(4P(\text{inside circle})) = \frac{16p(1-p)}{N} \tag{3.3}$$

or

$$\sigma_\pi = 4\sqrt{\frac{\frac{\pi}{4}(1-\frac{\pi}{4})}{N}} \propto \frac{1}{\sqrt{N}}. \tag{3.4}$$

3.1.2. A More Mathematical Description

As we have expressed the problem of the previous section in terms of areas, it is clear that it can also be formulated in terms of integrals. For this particular problem, we have:

$$\pi = \int_{-1}^{1} \mathrm{d}x \int_{-1}^{1} \mathrm{d}y \; \Theta(1-x^2-y^2), \tag{3.5}$$

where Θ is the Heaviside step function. More generally, consider the integral of some smooth function f over $[a,b] \subseteq \mathbb{R}$,[*]

$$I = \int_a^b \mathrm{d}x \; f(x). \tag{3.6}$$

Standard finite element methods for solving integrals of this type typically involve dividing the integration domain into N subintervals of length h and determining the weights w_i from e.g. a polynomial approximation. That is,

$$I \approx \sum_{i=1}^{N} w_i f(x_i). \tag{3.7}$$

The error σ in these sorts of methods is typically $\sigma \propto h^{-k} \propto N^{-k}$, where $k \in \mathbb{Z}_{>0}$. For a multi-dimensional integral, $\sigma \propto N^{-k/d}$, where d is the number of dimensions.[13]

In MC methods, $\forall i$ take $w_i = 1$ and $x_i \in [a,b]$ randomly sampled. That is,

$$I \approx \sum_{i=1}^{N} f(x_i). \tag{3.8}$$

If, for example, we choose x_i to be uniformly sampled, then the variance is

$$\sigma^2 = \left\langle \left(\frac{b-a}{N}\sum_{i=1}^{N} f_i\right)^2 \right\rangle - \left(\left\langle \frac{b-a}{N}\sum_{i=1}^{N} f_i \right\rangle\right)^2 \tag{3.9}$$

$$= \frac{(b-a)^2}{N}(\bar{f^2} - \bar{f}^2) \tag{3.10}$$

where $f_i := f(x_i)$ and x_i is a random number drawn, the angular brackets denote an average over all possible realisations, and the overbar represents an average of the function over the domain ($[a,b]$ in this discussion). i.e. the error in this method is proportional to the variance of f. Perhaps more interestingly, $\sigma \propto N^{-1/2}$, in line with the central limit

---

[*]We present the one-dimensional case for simplicity. The generalization to higher dimensions is straightforward.

theorem.[243] Comparing this error with standard quadrature, we see that MC integration is more efficient than an order-k algorithm when $d > 2k$. That is, although this particular MC example is naive, using simply a uniform distribution, it is still more efficient than standard methodologies for very large dimensions.

There exist several methods to reduce errors in MC methods.[244] Among the most important ones is importance sampling.[245] In the prevous example, it is clear that if significant contributions to the integral come from a small region of the domain, only a few points would be sampled by the MC algorithm there when using a uniform distribution. This would lead to large statistical errors. Mathematically, this is from the large variance of the function. In importance sampling, we sample from a distribution p which has roughly the same shape as f such that f/p is roughly constant over the integration domain. Of course, being a probability distribution, we require

$$p(x) > 0 \quad \forall x \tag{3.11}$$

and

$$\int \mathrm{d}x \ p(x) = 1. \tag{3.12}$$

The integral is

$$I = \int \mathrm{d}x \ f(x) = \int \mathrm{d}x \ \frac{f(x)}{p(x)} p(x). \tag{3.13}$$

Then, if we sample points according to p, we have

$$I \approx \frac{1}{N} \sum_{i=1}^{N} \frac{f(x_i)}{p(x_i)}, \tag{3.14}$$

where the naive MC method is recovered when $p(x) = 1/(b-a)$, i.e. the uniform distribution.

In this case, the variance in the result is

$$\sigma^2 = \left\langle \left(\frac{1}{N} \sum_{i=1}^{N} \frac{f(x_i)}{p(x_i)} \right)^2 \right\rangle - \left(\left\langle \frac{1}{N} \sum_{i=1}^{N} \frac{f(x_i)}{p(x_i)} \right\rangle \right)^2. \tag{3.15}$$

From this, we can see that if f/p is constant, the error vanishes, so it is important to choose a good p. If p is chosen poorly, the variance can be worsened, so this is a delicate problem.

In practice it is often difficult to estimate f, and a choice for p is problem-specific. However, when we have some approximation for its overall shape, importance sampling can be a powerful tool. Alternatively, another method known as adaptive Monte Carlo[246] seeks the most significant regions of f by random sampling, so that no a priori knowledge of the functional form is required.

3.1.3. Metropolis-Hastings Algorithm

A particularly successful method for generating random samples x_i from a probability distribution $\pi(x)$, where direct sampling of π is difficult, is the Metropolis-Hastings algorithm.[233,247] Often, $\pi(x)$ is known only up to a normalisation constant, $\pi(x) = \pi(x)/C$ where $C = \int \mathrm{d}x\,\pi(x)$ is intractible.

The Metropolis-Hastings algorithm is based on Markov chains,[248] and we construct such a Markov chain so that its stationary distribution is $\pi(x)$, which is to say if the Markov chain starts at $\pi(x)$ for step t, it is still $\pi(x)$ for step $t + 1$.

The probability of having a sequence of states $x_1, x_2, \ldots, x_N$ is

$$p(x_1, x_2, \ldots, x_N) = p(x_1)p(x_2|x_1)p(x_3|x_2)\cdots p(x_N|x_{N-1}). \tag{3.16}$$

where $p(x|x')$ is the transition probability from $x' \to x$. Then, the probability at step t to be in state x, $\pi_t(x)$ is given by

$$\pi_{t+1}(x) = \sum_{x' \in \Omega} \pi_t(x')p(x|x') \tag{3.17}$$

where Ω is the set of all possible states. This is called the master equation. In the stationary state, $\pi_t(x) = \pi(x)$, so

$$\pi(x) = \sum_{x' \in \Omega} \pi(x')p(x|x'). \tag{3.18}$$

Finding the general solution to this problem is nontrivial. However, a sufficient (but not necessary) condition is called detailed balance:

$$\pi(x)p(x|x') = \pi(x')p(x'|x). \tag{3.19}$$

This ensures that the probability of going from x to x' is the same as the probability of going from x' to x, which implies that the probability is stationary.

In order to actually construct the algorithm, we must introduce the trial step probability $\omega(x|x')$, and the acceptance probability $A(x|x')$. Then,

$$p(x|x') = \omega(x|x')A(x|x'). \tag{3.20}$$

$\omega(x|x'), A(x|x') \in [0, 1]$ for each pair x, x', and $\sum_{x'} \omega(x|x') = 1$. Furthermore, the original formulation of the algorithm required $\omega(x|x') = \omega(x'|x)$, which leads to (plugging into the master equation)

$$\frac{A(x|x')}{A(x'|x)} = \frac{\pi(x')}{\pi(x)}. \tag{3.21}$$

The algorithm proceeds in two stages: we propose a step $x' \to x$, and then accept or reject it. The probability of accepting is

$$A(x|x') = \min\left(1, \frac{\pi(x')}{\pi(x)}\right).$$ (3.22)

If we accept, we set the new state to x, otherwise we stay at x'.

In practice, on a computer, accepting is done by generating a uniform random number $r \in [0, 1)$ and accepting if $r < \frac{\pi(x')}{\pi(x)}$, and otherwise rejecting. Furthermore, we don't just have a single Markov chain, but a collection of so-called "walkers", that each perform their own Markov chain (for parallelisation). The integrand is then sampled at each position where the walkers reach.

One final note, is that since the current state of a Markov chain is dependent on the previous state, the Markov chain is not independent of itself. This is referred to as autocorrelation. One method to reduce this autocorrelation and give essentially independent samples is known as blocking.[249]

3.2. Variational (Quantum) Monte Carlo

As our first foray into QMC, we consider VMC. In section 1.3 we introduced the Variational Principle by parametrising the wave function and then finding the minimum of the expectation value of the energy occurring in the parametrisation space. If we have a large number of electrons and/or a large number of parameters, then the integrals involved in the evaluation of the energy will necessarily be high-dimensional. This is where MC comes in. For basic trial wave functions and small atoms like Hydrogen or Helium, direct integration may be possible, but as discussed in the previous section, this quickly becomes impossible for larger systems.

For a trial wave function Ψ (we omit the tilde from section 1.3), the expectation value of the energy is

$$\langle E \rangle = \frac{\langle \Psi | H | \Psi \rangle}{\langle \Psi | \Psi \rangle}$$ (3.23)

$$= \frac{\int \mathrm{d}^{3N}r \ \Psi^* H \Psi}{\int \mathrm{d}^{3N}r \ |\Psi|^2}$$ (3.24)

$$= \frac{\int \mathrm{d}^{3N}r \ \Psi^* \Psi \frac{H\Psi}{\Psi}}{\int \mathrm{d}^{3N}r \ |\Psi|^2}$$ (3.25)

$$= \frac{\int \mathrm{d}^{3N}r \ |\Psi|^2 E_L(\boldsymbol{r})}{\int \mathrm{d}^{3N}r \ |\Psi|^2},$$ (3.26)

where we have introduced the local energy, defined as

$$E_L(\boldsymbol{r}) := \frac{H\Psi}{\Psi}.$$ (3.27)

This rewriting of the integral is particularly suitable for evaluation using MC. Then we may vary the parameters and stop according to some minimisation algorithm. Notice that the more strongly Ψ resembles the exact wave function, the less strongly E_L varies with $\boldsymbol{r}$. In particular, if Ψ is equal to an exact eigenstate, then E_L is constant. Therefore, an alternative objective function to the energy expectation of equation 3.23 is the variance.[250–252]

Defining

$$p(\boldsymbol{r}) = \frac{|\Psi|^2}{\int \mathrm{d}^{3N}r \ |\Psi|^2}, \tag{3.28}$$

we may write equation 3.23 as

$$\langle E \rangle = \int \mathrm{d}^{3N}r \ p(\boldsymbol{r}) E_L(\boldsymbol{r}). \tag{3.29}$$

This form of the energy is amenable to the Metropolis-Hastings approach outlined in section 3.1.3. Recall also that p need not be normalised when using the Metropolis-Hastings algorithm for sampling.

In principle, the algorithm is doable with just a single walker, but in practice we reduce the statistical error by using many. The algorithm then becomes

```
Initialise  N  walkers  in  random  positions .
until  convergence  criterion  met*
   for  each  walker  at  position  r
     sample  the  local  energy  E_L  at  r
     propose  a  new  position  r'  with  probably  p = |Ψ(r')/Ψ(r)|²
     if  r'  is  accepted
        set  r  to  r'
```

The energy is then calculated as the expectation value of the local energy, averaged over the samples generated in this procedure. Steps at the beginning (before equilibrium) are discarded in a process called equilibration,[253] and a blocking procedure should be employed. The decision to stop is generally based on compute time and/or precision required.

3.2.1. Trial Wave Functions

While VMC is a powerful tool, the quality of the solution is constrained by the quality of the trial wave function. Moreover, the evaluation of the trial wave function is expensive, and we therefore want a form that is easy to evaluate.

---

*The convergence criterion in VMC is typically just to sample a predefined number of configurations.

The form of the trial wave function is typically chosen based on a Jastrow factor, as already introduced in section 2.5. Generally,[174]

$$\Psi_{\text{trial}} = e^J \Phi. \tag{3.30}$$

For computational efficiency, the Jastrow factor typically only retains one- and two-body terms,[230]

$$J = \sum_i \chi(\boldsymbol{x}_i) - \frac{1}{2} \sum_{i \neq j} u(\boldsymbol{x}_i, \boldsymbol{x}_j), \tag{3.31}$$

where χ describes electron-nuclear correlation and u describes two-electron correlation (including e.g. electron-electron-nuclear correlation). There exist a number of more specific forms for J,[225] though one unifying principle is for them to adhere to expected short-range (cusp conditions) and long-range ($1/r$) behaviour.

Φ is typically chosen to be a single Slater determinant.[230,254–257] In particular, by choosing the HF determinant, typically only a small portion of the energy (that is, the correlation energy) is left for the Jastrow factor to describe, and typically even simple Jastrow factors can do this.[230] However, sometimes this is not enough, for example, when the HF determinant is not sufficient to describe all symmetries of the true wave function.[258–260]

3.3. Diffusion Monte Carlo

Diffusion Monte Carlo (DMC) starts by performing a Wick rotation (that is, substitute $t \to i\tau$) on the time-dependent Schrödinger equation, equation 1.1. We then get the imaginary-time Schrödinger equation

$$\frac{\partial}{\partial \tau} \Psi(\boldsymbol{r}, \tau) = -\hat{H} \Psi(\boldsymbol{r}, \tau). \tag{3.32}$$

This equation is a diffusion (or heat) equation, and Ψ may be interpreted as the density distribution for a large number of independent particles* (walkers).[261] More specifically, the kinetic energy term, being second order, describes the movement of the walkers whereas the potential energy term describes generation or annihilation of walkers. The solution of this equation is

$$\Psi(\boldsymbol{r}, \tau) = e^{-(\hat{H} - S)\tau} \Psi(\boldsymbol{r}, 0), \tag{3.33}$$

where we have introduced the energy shift S which is adapted throughout the simulation every n steps to keep the number of walkers constant, and to estimate the ground state energy.[262] In particular, if we expand $\Psi(\boldsymbol{r}, 0)$ in terms of the eigenfunctions of $\hat{H}$, $\{\psi_i\}$,

---

*Note that the particles, henceforth referred to as walkers, are not the particles being described by the Hamiltonian. One walker represents a configuration of all the particles described by the Hamiltonian.

labeled in such a way that $i > j \implies E_i > E_j$, we get

$$\Psi(\boldsymbol{r}, \tau) = \sum_i a_i \mathrm{e}^{-(\hat{E}_i - S)\tau} \psi(\boldsymbol{r}). \tag{3.34}$$

Here, a_i is the coefficient of the ith eigenfunction, and $\hat{E}_i$ is the corresponding eigenvalue. After the excited-state solutions have decayed, the only possible value of S that stabilises the simulation is $S = E_0$. It is also clear from equation 3.34 that when $S = E_0$, all excited states decay exponentially to zero, and we are left with the ground state for large τ.

We may also express Ψ in terms of its Green's function, in order to get a form that expresses a small time step,

$$\Psi(\boldsymbol{r}, \tau + \Delta\tau) = \int \mathrm{d}^{3N} r' \; G(\boldsymbol{r}, \boldsymbol{r}'; \Delta\tau) \Psi(\boldsymbol{r}, \tau), \tag{3.35}$$

where the Green's function

$$G(\boldsymbol{r}, \boldsymbol{r}'; \Delta\tau) = \langle r | \, \mathrm{e}^{-\Delta\tau(H-S)} \, | r' \rangle \tag{3.36}$$

also obeys the imaginary-time Schrödinger equation, but with the initial condition

$$G(\boldsymbol{r}, \boldsymbol{r}'; 0) = \delta(\boldsymbol{r} - \boldsymbol{r}'). \tag{3.37}$$

For small time steps, G may be approximated by the Lie-Trotter-Suzuki decomposition,[263–265] i.e. for operators A, A_k,

$$A = \sum_k A_k \implies \mathrm{e}^{-\alpha A} \approx \left(\prod_k e^{-\alpha A_k/M} \right)^M \tag{3.38}$$

for some value α and for large M. From this, we can get the Green's function,

$$G(\boldsymbol{r}, \boldsymbol{r}'; \Delta\tau) = \mathrm{e}^{-\Delta\tau(V(r)-S)} \frac{1}{\sqrt{2\pi\Delta\tau}} \mathrm{e}^{-(r-r')^2/(2\pi\Delta\tau)} + \mathcal{O}(\Delta\tau^2), \tag{3.39}$$

which describes our time evolution.

In DMC, walkers diffuse through configuration space in two stages: diffusion and branching. In the diffusion step, walkers are moved according to the kinetic terms in the Green's function. In the branching step, unfavourable walkers are removed and favourable ones produce new walkers.

For the energy shift estimate, we wish to choose a value such that the overall population does not change too much. If the walkers proliferate too much, then the computational cost becomes infeasible, whereas if too many walkers die, we are left with poor statistics. One way to choose S is by keeping track of the population of walkers N_w for a target

populaton N_w^{target} and adjust,

$$S = S_{\text{old}} + \alpha \ln \left(\frac{N_w^{\text{target}}}{N_w} \right) \tag{3.40}$$

for small α.

More concretely, in pseudocode, the basic DMC algorithm is:

```
Initialise N walkers in random positions.
until convergence criterion met
    for each walker at position r
        shift position r to r' according to
        ↪transition probability G
        evaluate q = e^{-Δτ(V(r')-S)}
        walker survives with probability q
        if q > 1
            create 1 - q new walkers
                ↪(with stochastic rounding)
    update S.
```

Notice that in contrast to VMC, DMC does not rely on a trial wave function. There is also a discretisation error incurred by the operator decomposition, though this can be handled via an application of the Metropolis-Hastings alogirthm.[231] However, in practice a trial wave function is actually necessary for fermions in order to improve the statistics. The walker distribution is positive, but for fermions this is not the case due to their exchange statistics. Since in electronic structure, we focus on fermions (electrons), this is an important feature. In order to still practically use DMC,* an approximation called the fixed node approximation is made, where the nodal structure (i.e. the roots of Ψ) of the distribution is determined by a trial function. Thus, the accuracy of DMC is still limited by the quality of the trial wave function (or more precisely, by its nodal surface). In a Slater-Jastrow-type trial wave function like equation 3.30, only the Slater-component will affect the nodal surface (and hence the accuracy). Thus, the choice of orbitals as well as the number of determinants can be important considerations.

3.4. Full Configuration Interaction Quantum Monte Carlo

The main post-Hartree-Fock method used throughout this dissertation is the FCIQMC algorithm.[266] For this, there exist a few different implementations,[267–270] but we focus on NECI.[267] As the name implies, FCIQMC combines concepts from FCI (discussed in section 1.6.1) and QMC concepts already discussed in this chapter. In particular, it implements FCI, retaining many of its desirable properties such as size consistency and accuracy while

---

*Other simple solutions such as assigning a sign to each walker leads to large errors.

being able to handle much larger systems thanks to its stochastic framework (however, the scaling remains exponential, like FCI). Like FCI, it is non-perturbative, and so can handle multi-reference problems, and unlike other QMC methods, it does not rely on a trial wave function. It can also be used as a solver in place of FCI, like in a MCSCF calculation.[271,272]

3.4.1. Main Concepts

Like DMC, FCIQMC is based on the imaginary-time Schrödinger equation, equation 3.32, repeated here:

$$\frac{\partial}{\partial \tau}\Psi(\boldsymbol{r},\tau) = -\hat{H}\Psi(\boldsymbol{r},\tau). \tag{3.41}$$

The solution to this equation is

$$\Psi(\boldsymbol{r},\tau) = \mathrm{e}^{-\tau\hat{H}}\Psi(\boldsymbol{r},0). \tag{3.42}$$

Expanding $\mathrm{e}^{-\tau\hat{H}}$ to first order in small imaginary time step $\Delta\tau$ gives us the FCIQMC projector

$$P = 1 - \Delta\tau\hat{H} \rightarrow 1 - \Delta\tau(\hat{H} - S), \tag{3.43}$$

where we have introduced an arbitrary shift S in analogy to DMC. Notice, in particular, that if $S = E_0$ the ground state energy, all other states decay away. Hence, repeated application of P while adjusting S to get as close to E_0 as possible will give us the ground state, as long as our initial state has nonzero overlap with the ground state.

Like FCI, the wave function is expanded in some basis. Throughout this dissertation, SDs are used, though configuration state functions (CSFs) are also possible.[273] The basis will be kept constant, and the coefficients may change with each time step,

$$|\Psi(\tau)\rangle = \sum_i c_i(\tau) |D_i\rangle . \tag{3.44}$$

Like in DMC, we adjust S to approximate the ground state energy E_0, as $S = E_0$ would project out all states other than the ground state, as long as $\langle\Psi(\tau = 0)|\Psi_0\rangle \neq 0$ where $|\Psi_0\rangle$ is the ground state. Similarly, S affects the walker population. In particular, for walker population N_w, if $S > E_0$ then the population grows, and if $S < E_0$ it decays. If we have a target population N_w^{target} then we keep $S > E_0$ (typically, just by setting $S = 0$) and "turn on" the shift once N_w^{target} is reached. However, unlike DMC, FCIQMC walkers can have either negative or positive weight. These weights correspond to the c_i, and the walkers "diffuse" across SD (or CSF) space. The walker population is reported in terms of absolute values, i.e.

$$N_w = \sum_{i\in\{SDs\}} |N_i|, \tag{3.45}$$

where we sum over the walkers across all SDs. Once in variable-shift mode, in practice S is updated every m cycles:

$$S(\tau + m\Delta\tau) = S(\tau) - \frac{\zeta}{m\Delta\tau} \ln\left(\frac{N_w(\tau + m\Delta\tau)}{N_w(\tau)}\right), \tag{3.46}$$

where ζ is a damping parameter.

The key advantage of FCIQMC over ordinary FCI is no longer needing to store the full CI vector, making use of sparsity. Even for relatively large and strongly-correlated systems, only a small subspace of the entire Hilbert space needs to be sampled in FCIQMC, whereas in FCI, the full CI vector must be stored (or if done via dense diagonalisation, the entire H matrix, which is unrealistic for even small systems).

3.4.2. Basic Algorithm

FCIQMC is typically run in three steps: (a) **spawning**, (b) **death/cloning**, and (c) **annihlation**. Each step can be recognised in the application of the projector operator of equation 3.43 onto some state $|\Psi\rangle = \sum_i c_i |D_i\rangle$,

$$P|\Psi\rangle = |\Psi\rangle \underset{\substack{\uparrow \\ (c)}}{-} \underbrace{\sum_i \sum_{j\neq i} \Delta\tau H_{ji} c_i |D_j\rangle}_{(a)} \underset{\substack{\uparrow \\ (c)}}{-} \underbrace{\sum_i (\Delta\tau(H_{ii} - S))c_i |D_i\rangle}_{(b)}, \tag{3.47}$$

where the right-hand side of the equation is the new wave function, $|\Psi(\tau + \Delta\tau)\rangle$. Or, in terms of the coefficients (relabelling $i \leftrightarrow j$ in the (a) terms),

$$c_i(\tau + \Delta\tau) = c_i(\tau) - \Delta\tau(H_{ii} - S)c_i(\tau) - \Delta\tau \sum_{j\neq i} H_{ij} c_j(\tau). \tag{3.48}$$

In practice, we start the algorithm with a population of walkers on a reference determinant (typically, but not always, the Hartree-Fock determinant), and calculate the correlation energy via FCIQMC. Instead of H in the working equations we would use the matrix

$$K_{ij} = H_{ij} - E_{\text{ref}}\delta_{ij} \tag{3.49}$$

where $E_{\text{ref}} := \langle D_{\text{ref}}| H |D_{\text{ref}}\rangle$ is the energy of the reference determinant. However, FCIQMC is capable of adapting to a new reference determinant whenever another determinant exceeds the population of the reference determinant, and is not single-reference like canonical coupled cluster, for example. This shift is primarily a convenience.

The basic FCIQMC algorithm can be summarised by the following pseudocode (with further details below). Note that there are many extensions to the basic algorithm, beyond what is discussed in this dissertation.

```
Initialise N_start walkers on the reference determinant.
until stop criterion met
    for each occupied determinant D_i
        select a random determinant based on p_gen
        spawn p_s(j|i) walkers on D_j
    for each occupied determinant D_i
        remove p_d (death/cloning) walkers from D_i
    for each occupied determinant D_i
        annihilate opposite signed walkers
    calculate the projected energy
    if  ∑|N_i| ≥ N_w
        update S.
```

Spawning

The first step, spawning, is realised by the application of the off-diagonal matrix elements. In this step, each walker on SD D_i may spawn a new set of walkers to connected* SD D_j. As we are sampling the sum $\Delta\tau \sum_{j \neq i} H_{ij} c_j(\tau)$, as in the Metropolis-Hastings algorithm, it is beneficial (in terms of efficiency) to sample more where H_{ij} is larger. That is, the probability of spawning on D_j from D_i, denoted $p_{\text{gen}}(j|i)$ (the generation probability) should be proportional to H_{ij}. Finally, the spawning probability is

$$p_s(j|i) = \Delta\tau \frac{|H_{ij}|}{p_{\text{gen}}(j|i)}. \tag{3.50}$$

Thus, at each FCIQMC iteration, every determinant D_i spawns $N_i p_s(j|i)$ new walkers on a determinant D_j selected by an excitation generation algorithm.[274–276] Here, N_i is the (signed) number of walkers on D_i. If $H_{ij} < 0$ then the child walkers are the same sign as the parent; otherwise, they have opposite sign. The original formulation of the algorithm used integer number of walkers, and spawned a walker with probability $p_s(j|i)$ per walker on N_i, but it is more efficient to simply use noninteger values and spawn $p_s(j|i)$ walkers.

Death/Cloning

The death/cloning step is characterised by the diagonal matrix elements. In this step, p_d walkers are removed from the simulation† where

$$p_d = \Delta\tau(H_{ii} - S). \tag{3.51}$$

---

*D_i and D_j are said to be connected if $\langle D_i | H | D_j \rangle \neq 0$.

†The original formulation removed (or cloned) walkers with a probability of p_d.

Notice that if p_d is the same sign as the walker, then the number of walkers on D_i increases (hence "cloning", as opposed to "death"). "Cloning" is especially common in the growth stage of the population, where the population is growing exponentially towards the target N_w.

Annihilation

After spawning and death/cloning, we may have walkers of opposite sign on the same determinant. This is dealt with in the "annihilation" step of the algorithm, realised by combining the terms in equation 3.47. This step, while simple in principle, amounts to cancelling opposite walkers. Determinants without walkers left on them are then removed as well, to keep the storage of $|\Psi\rangle$ sparse. However, so far the steps have all been embarrassingly parallel, but in this step we need to communicate in order to locate determinants on the same determinant, so this step is critical for parallelisation. Nevertheless, efficient implementations such as in `NECI` is able to scale efficiently to over 40000 cores.[277]

3.4.3. Energy Estimators

The shift S is updated to approximate E_0, so we already have an estimator for the energy. However, since S is used to also control the population, this estimator can be noisy.[266]

A more commonly used energy estimator is the projected energy,

$$E_{\text{proj}} = \frac{\langle D_{\text{ref}}| \hat{H} |\Psi\rangle}{\langle D_{\text{ref}}|\Psi\rangle}, \tag{3.52}$$

where we project onto the reference (most occupied) determinant.*

In principle, any wave function could be used in place of the trial determinant in equation 3.52 (so long as it has nonzero overlap with $|\Psi\rangle$). The choice of the reference determinant is for ease of computation, and because the larger the overlap with $|\Psi\rangle$, the smaller the error in the stochastic sampling. Since the reference determinant (by definition) has the largest number of walkers on it, it is the sensible choice for a single-determinant projected energy.

Another energy estimator is the "trial energy",[279,280] which is just the projected energy on a multi-determinant wave function instead of $|D_{\text{ref}}\rangle$. In this method, after some time in variable shift mode, a "trial space" is constructed based on the top N_T determinants. The Hamiltonian spanned within the trial space is diagonalised exactly (sparsely or densely), and the eigenvector $|\Psi_{\text{trial}}\rangle$ is stored. Then, the trial projected energy can be calculated as

$$E_{\text{trial}} = \frac{\langle \Psi_{\text{trial}}| \hat{H} |\Psi\rangle}{\langle \Psi_{\text{trial}}|\Psi\rangle}. \tag{3.53}$$

---

*While it may be tempting to instead use the energy estimator $\langle\Psi| \hat{H} |\Psi\rangle / \langle\Psi|\Psi\rangle$, this comes with a prohibitive cost, and has a bias.[266,278]

Note that while this is expected to reduce the error in the energy estimator, it also increases the cost by a factor of N_T.

An example of these energy estimators is shown in figure 3.2. As in other QMC methods, these energy estimators have autocorrelation, so in reporting the values, a blocking analysis is required.[249]

3.4.4. Annihilation Plateaus

FCIQMC has its own kind of sign problem in the form of resolving the correct relative sign structure of the sampled states in the CI vector.[281] This manifests in the form of a so-called "annihilation plateau", illustrated in figure 3.3.

In this stage of the simulation, walkers are being spawned with incoherent signs, causing a massive amount of walker annihilation. This causes the population to stay relatively constant, until the sign structure is finally resolved and the population continues to grow again. The height of the plateau (i.e. the constant number of walkers where the simulation stagnates) quantifies the difficulty of the sign problem. In particular, for certain large systems, it can be prohibitively expensive to overcome. If we are below the annihilation plateau, then the original formulation of FCIQMC is not reliable, as the correct sign structure is not yet resolved.

3.4.5. The Initiator Approximation

Since incoherent walker proliferation is a major source of the annihilation plateau, a sensible strategy to overcome this is by forcing growth to be coherent. This is achieved by the initiator approximation.[282] An FCIQMC calculation with the initiator approximation is sometimes referred to as i-FCIQMC; however, since the approximation is so useful and ubiqituous, we simply referred to it as FCIQMC. Indeed, in this dissertation all FCIQMC calculations used the initiator approximation, unless otherwise specified.

The initiator approximation allows only those determinants (dubbed "initiators") with at least N_{thresh} walkers to spawn new walkers on unoccupied determinants. If the population is greater than this threshold, the assumption is that the (relative) sign of that state is correct, and hence the spawning event should produce the correct sign.

Since the approximation truncates spawning, the growth phase of the simulation is slower. However, the annihilation plateau disappears. Note that as $N_w \rightarrow \infty$, the initiator approximation becomes exact, as all sampled states are initiators. While there is a small bias introduced by the initiator approximation, the efficiency is greatly improved and the range of potential applications is greatly increased.

3.4.6. Reduced Density Matrix Sampling

As described in section 1.4.2, reduced density matrices are a useful tool for understanding the properties of a system. These can be sampled in FCIQMC by considering the RDMs
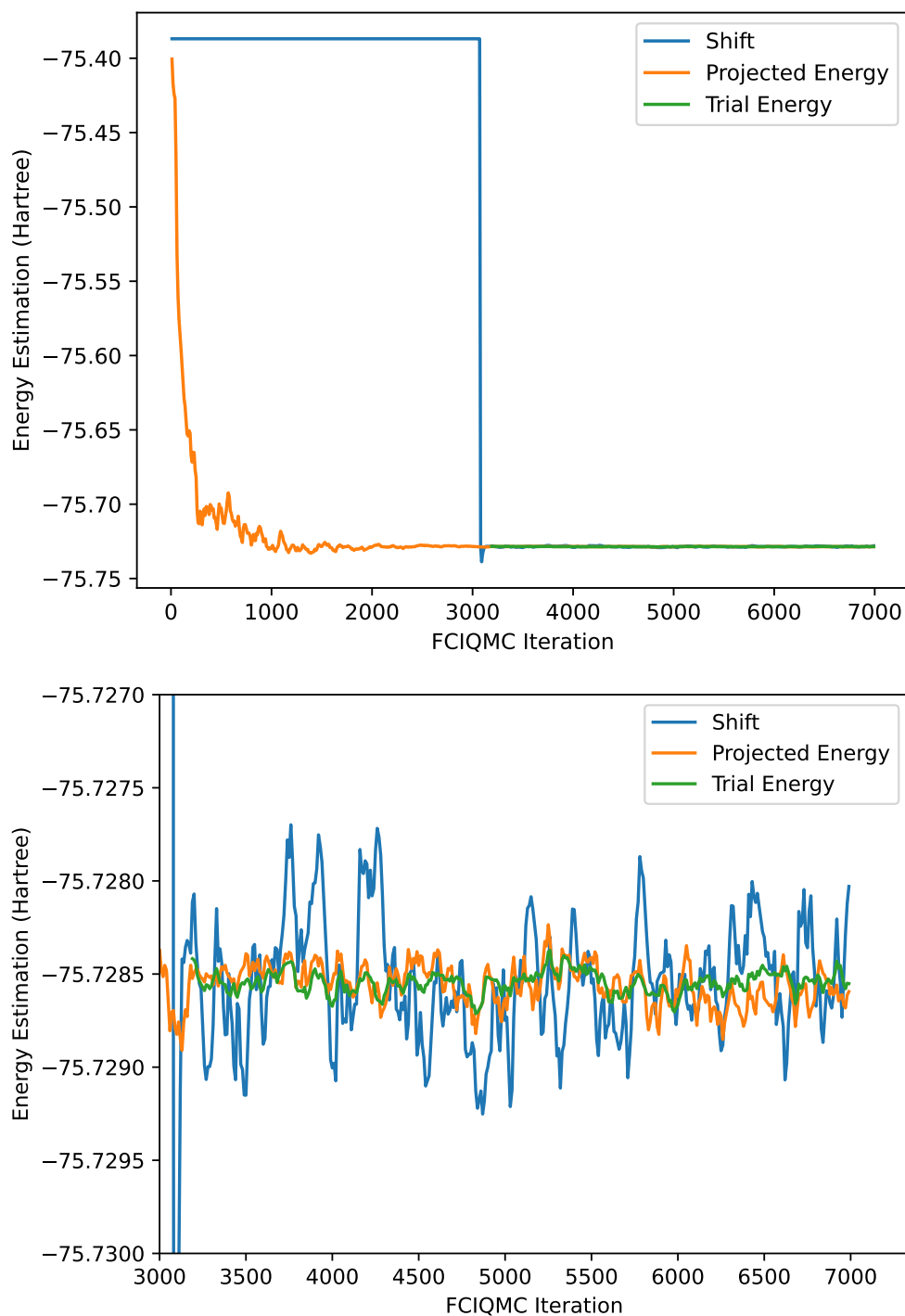
Figure 3.2. An example of the energy estimators used in an FCIQMC simulation. The shift is the most noisy, whereas the trial energy is the least noisy, but only available in variable shift mode and carries the highest cost. In this case, the trial energy is not a substantial improvement from the reference-projected energy. This calculation was done on the C_2 molecule with the cc-pVDZ basis, at equilibrium geometry, 1.2425 Å.
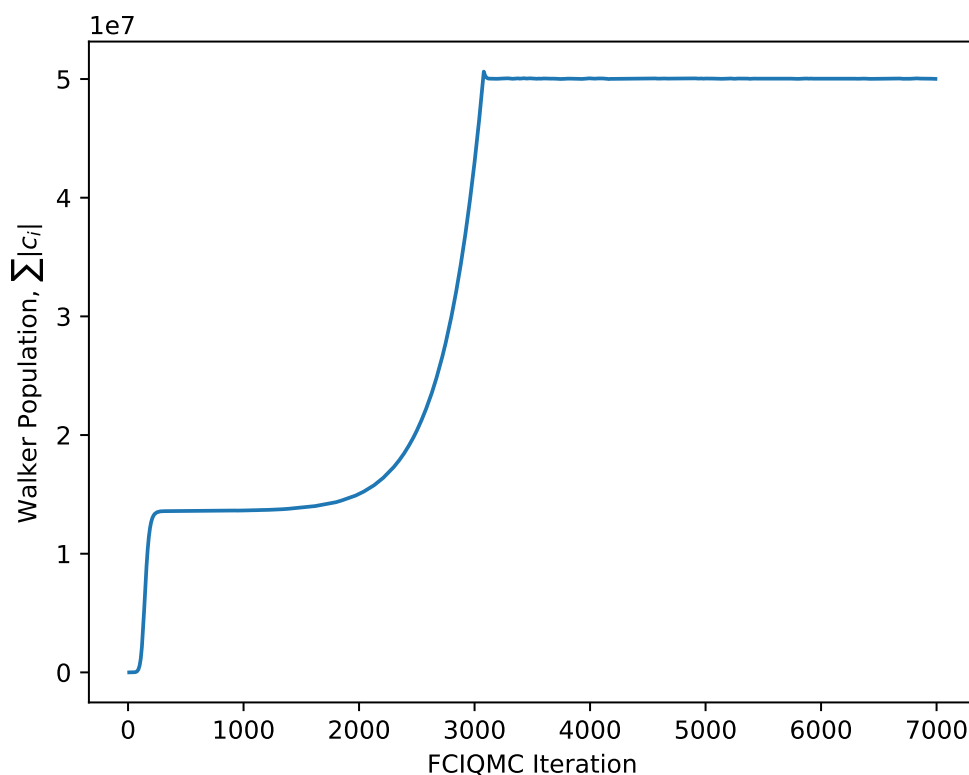
Figure 3.3. An example of an annihilation plateau in the context of a FCIQMC simulation. Despite not yet reaching the target population, the total population is roughly constant before increasing again. This stage is referred to as the annihilation plateau, and the higher the plateau, the more difficult the sign problem is. This calculation was done on the C_2 molecule with the cc-pVDZ basis, at equilibrium geometry, 1.2425 Å. In this particular case, the plateau appears around 1.36×10^7 walkers, and the target number is 5×10^7.

in the same basis.[283] The 1RDM is given by

$$\gamma_{pq} = \sum_{ij} c_i c_j \langle D_i | a_p^\dagger a_q | D_j \rangle \tag{3.54}$$

and the 2RDM by

$$\Gamma_{pqrs} = \sum_{ijkl} c_i c_j \langle D_i | a_p^\dagger a_q^\dagger a_s a_r | D_j \rangle . \tag{3.55}$$

In FCIQMC, we have the CI coefficients c_i and c_j by the walker distributions. However, using the same simulation for both c_i and c_j introduces a bias, hence to stochastically sample RDMs, we use an independent simultaneous simulation (referred to as a replica). One simulation samples c_i whereas the other samples c_j. Since the statistical errors are independent, the overall error is reduced. Unfortunately this requires twice the computing power, but thankfully since the calculations are independent, they are trivially parallelisable.

Note also that by inserting the Hamiltonian (energy) operator into equation 1.21 and using the replica trick to obtain the RDMs, we can also obtain yet another energy measure in FCIQMC.

3.4.7. Combining TC with Modern Electronic Structure

The FCIQMC algorithm may also be combined with the TC method described in section 2.5.[205] Consider solving for the eigenstates Φ of the TC Hamiltonian $\hat{H}_{\text{TC}}$ using the imaginary-time Schrödinger equation,

$$-\frac{\partial}{\partial \tau} \Phi = (\hat{H}_{\text{TC}} - S)\Phi. \tag{3.56}$$

Since $\hat{H}_{\text{TC}}$ and $\hat{H}$ are isospectral, we have stationary states for the same values as the non-TC method. That is, we can control the walker population by setting the shift to the ground state energy, $S = E_0$.

The state Φ is described by a linear combination of SDs, as in non-TC-FCIQMC,

$$|\Phi\rangle = \sum_i c_i |D_i\rangle . \tag{3.57}$$

However, since $\hat{H}_{\text{TC}}$ is non-Hermitian, the coefficients $\tilde{c}_i$ of the left-eigenvector with the same energy are not necessarily the same,

$$\langle \Phi | = \sum_i \tilde{c}_i \langle D_i | , \tag{3.58}$$

and we must be careful when considering matrix connections $\langle D_i | \hat{H}_{\text{TC}} | D_j \rangle$ and $\langle D_j | \hat{H}_{\text{TC}} | D_i \rangle$. In particular, the probability of spawning a walker from $|D_i\rangle$ to $|D_j\rangle$ may not be the same

as the probability of spawning from $|D_j\rangle$ to $|D_i\rangle$. Furthermore, when considering replicas such as in RDM sampling, the replica needs to be of the adjoint operator, $\hat{H}_{\text{TC}}^\dagger$, in order to target $\tilde{c}_i$. This is accomplished by a simple transform of the Jastrow factor, $J \to -J$, though this may cause practical complications. Otherwise, FCIQMC may be extended to TC by simply applying the method directly to $\hat{H}_{\text{TC}}$. An appropriate transformation from $\hat{H}$ is therefore necessary beforehand.

Similar arguments may be made for applying TC to CC methods[213,214] and DMRG,[189] and TC variants of these methods have already been developed and successfully applied. In addition to TC-FCIQMC, this also continues to be an active area of research.

Optimising Jastrow Factors for the Transcorrelated Method

This chapter is based in large part on the following paper, and most of the following discussion can already be found there:

Haupt, J. P.; Hosseini, S. M.; López Ríos, P.; Dobrautz, W.; Cohen, A.; Alavi, A. "Optimizing Jastrow Factors for the Transcorrelated Method". The Journal of Chemical Physics **2023**, *158*, 224105

Images have been reused from this paper (with permission).

4.1. Introduction

In this chapter, we investigate the use of flexible Jastrow factors and a novel optimisation strategy for use in TC as introduced in section 2.5. As a brief recapitulation, the TC method amounts to a similarity transformation of the Hamiltonian $\hat{H}$, $\hat{H}_{\mathrm{TC}} = \mathrm{e}^{-J} \hat{H} \mathrm{e}^{J}$. However, as this is a non-unitary trasformation, methods used to solve $\hat{H}_{\mathrm{TC}}$ are in general not variational, and hence we are not guaranteed to converge to the CBS limit from above. It is therefore important to choose J wisely, as otherwise the method may be highly non-variational, and we may suffer from poor error cancellation.

As an illustration of the method, we compute the all-electron atomisation energies for the challenging first-row molecules C_2, CN, N_2 and O_2 and find that TC-FCIQMC (that is, FCIQMC performed on a transcorrelated Hamiltonian) yields chemically accurate results using only a cc-pVTZ basis, which requires a much larger cc-pV5Z basis for non-TC.

4.2. Computational Details

We compute the ground-state energies of the all-electron C, N, and O atoms, as well as that for the C_2, CN, N_2 and O_2 molecules at their equilibrium geometries,[285–287] listed in table 4.1. TC- and non-TC-FCIQMC calculations used HF orbitals (restricted open-shell in the case of open-shell systems) expanded in the standard cc-pVXZ family of basis sets.[288]

The quality of the energy differences is assessed using the atomisation energies of these molecules. In order to determine if our methodology yields chemically-accurate, i.e. within

Table 4.1. Electronic ground states and equilibrium bond lengths used for the molecules considered in this work, following reference 285.

System	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

an error of 1 kcal/mol $\approx$ 1.6 mHa, we also keep each individual error to be well within this threshold. We expect a total bias in our resulting relative energies of not more than 0.5 mHa.

For all our calculations, we generate our orbitals and integration grids using `pyscf`,[289] optimise the Jastrow factors using the `CASINO` continuum QMC package,[253] compute TC matrix elements using the `tchint` library, for which more details are presented in Appendix A, and perform (TC-)FCIQMC calculations using the `NECI` package.[267] FCIQMC energies reported are the standard HF-projected energies.

(TC-)FCIQMC values presented here were produced using a walker-number extrapolation scheme presented in another dissertation.[290]

4.3. Jastrow Factor

In continuum quantum Monte Carlo methods, the Jastrow factor for a molecule is typically expressed as the sum of electron-electron, electron-nucleus, and electron-electron-nucleus terms,[*]

$$J = \sum_{i<j}^{N} v(r_{ij}) + \sum_{i}^{N}\sum_{I}^{N_A} \chi(r_{iI}) + \sum_{i<j}^{N}\sum_{I}^{N_A} f(r_{ij}, r_{iI}, r_{jI}), \tag{4.1}$$

where N_A is the number of nuclei, N the number of electrons, and each of u, χ, and f are expressed as natural power expansions.[224] That is,

$$v(r_{ij}) = t(r_{ij}, L_v) \sum_k a_k r_{ij}^k, \tag{4.2}$$

$$\chi(r_{iI}) = t(r_{iI}, L_\chi) \sum_k b_k r_{iI}^k, \tag{4.3}$$

$$f(r_{ij}, r_i, r_j) = t(r_{iI}, L_f) t(r_{jI}, L_f) \sum_{k,l,m} c_{klm} r_{ij}^k r_{iI}^l r_{jI}^m, \tag{4.4}$$

---

[*]Of course, these are not all the possible terms. We may, for example, also choose to include electron-nucleus-nucleus terms.

where $\{a_k\}$, $\{b_k\}$, and $\{c_{klm}\}$ are linear parameters, L_v, L_χ, and L_f are cut-off lengths, $t(r, L) = (1 - r/L)^3 \Theta(r - L)$ is a cut-off function, and $\Theta(r - L)$ is the Heaviside step function.

As described in chapter 2, accurately describing the (electron-electron and electron-nucleus) Kato cusp conditions[67] substantially improves the accuracy of our method. Also, as described in chapter 3, VMC and DMC methods sample electronic configurations $\{\boldsymbol{R}\}$ from a probability distribution based on an analytical trial wave function $\tilde{\Psi}_\mathrm{T}(\boldsymbol{R})$ to produce a variational estimate of the total energy as an average of the local energy, $E_\mathrm{L}(\boldsymbol{R}) = \tilde{\Psi}_\mathrm{T}^{-1}(\boldsymbol{R}) \hat{H}(\boldsymbol{R}) \tilde{\Psi}_\mathrm{T}(\boldsymbol{R})$ over the sampled configurations. In the case of VMC, accurate description of the electron-electron and electron-nucleus Kato cusp conditions suppresses extreme outliers in the local energy sampling, allowing meaningful wave function parameters.

The most obvious way to enforce the electron-electron and electron-nucleus cusp conditions is by enforcing them in the form of the Jastrow factor through the relevant terms, namely v (equation 4.2) for the electron-electron cusp, and χ (equation 4.3) for the electron-nucleus cusp. However, in the context of continuum QMC, it has been found to be better[224,253,291] to enforce the electron-nucleus cusp by modifying the $l = 0$ (s-type) component of the cuspless molecular orbitals, $\phi(r)$, such that they exhibit a cusp.

Since we are interested in performing a post-Hartree-Fock calculation on $\hat{H}_\mathrm{TC}$, such as FCIQMC, it is preferable to use unmodified molecular orbitals from standard basis sets during the optimisation process. If we optimise the Jastrow factor in VMC in the presence of cusp-corrected orbitals and then use them in TC-FCIQMC without the cusp-corrected orbitals, the Jastrow factor would be sub-optimal for the Hamiltonian, by construction.

Instead, we recast the cusp-correction scheme of Ref. 291 as an electron-nucleus Jastrow factor term, called Λ, to be added (rather than replacing) the χ term of equation 4.3. We construct this term as

$$\Lambda(r) = \left[\ln \tilde{\phi}(r) - \ln \phi(r) \right] \Theta(r - r_c), \tag{4.5}$$

where, adopting the notation of Ref. 291, r_c is a cutoff radius, $\phi(r)$ is the s-type component of the target orbital, and $\tilde{\phi}(r)$ is its cusp-corrected counterpart,

$$\tilde{\phi}(r) = \mathrm{e}^{\sum_{l=0}^{4} \alpha_l r^l} + C \quad , \quad r < r_c. \tag{4.6}$$

Here, $\{\alpha_l\}$ are parameters determining the shape of the corrected orbital and the shift C is only set to a non-zero value in the presence of nodes of $\phi(r)$ near the nucleus. More precisely, the shift C is chosen such that $\tilde{\phi}(r_c) - C$ is of one sign within the radius r_c. This is necessary since we wish to impose an exponential correction, which is necessarily of one sign.

Following Ref. 291, we impose the cusp condition at $r = r_c$, as well as twice continuous differentiability at $r = r_c$. This leaves only α_0 and r_c as free parameters from equations

4.5 and 4.6. r_c is chosen to be small but within the same sign, as described above, while α_0 is determined by enforcing smoothness for the so-called "effective one-electron local energy",

$$E_L^s(r) := \tilde{\phi}(r)^{-1} \left[-\frac{1}{2}\nabla^2 - \frac{Z_{\text{eff}}}{r} \right] \tilde{\phi}(r). \tag{4.7}$$

Here, the effective nuclear charge Z_{eff},

$$Z_{\text{eff}} = Z \left(1 + \frac{\eta(0)}{\tilde{\phi}(0)} \right) \tag{4.8}$$

ensures that E_L^s is finite at the origin, and is derived from the cusp condition. η is the rest of the orbital, leftover from removing the s-type component.

Figure 4.1 illustrates the effect of using a Λ term in practice.
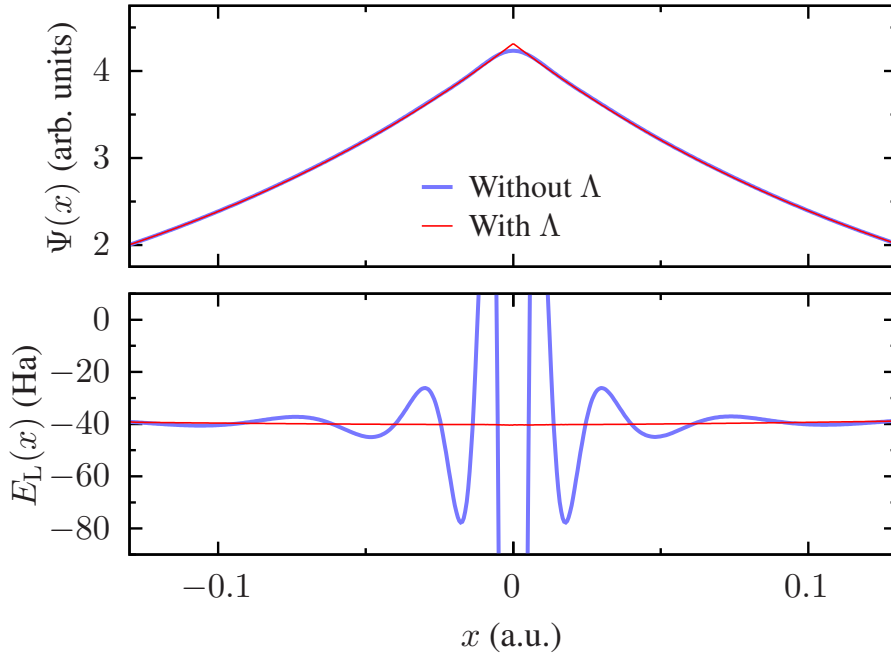
Figure 4.1. HF wave function value and local energy as a function of the x coordinate of an electron in a carbon atom as it crosses the nucleus at $x = 0$, both with and without the Λ cusp-correcting Jastrow factor term. This is in the cc-pVDZ basis.

For the calculations in this chapter, we use a total of 44 optimisable Jastrow factor parameters for the atoms and homonuclear dimers, and 88 parameters for CN. We keep the L_v, L_χ and L_f cutoff lengths fixed at 4.5, 4, and 4 Bohr, for simplicity.

4.4. Optimisation Strategy

We optimise J using VMC. VMC provides a variational framework in which parameters $\boldsymbol{\alpha}$ present in a trial wave function Ψ_T can be optimised. In this chapter, $|\Psi_T\rangle = e^{J(\boldsymbol{\alpha})} |D_{\text{HF}}\rangle$.

In continuum QMC, wave function optimisation is usually carried out using a correlated-sampling approach in which a set of n_{opt} electronic real-space configurations $\{\boldsymbol{R}_i\}_{i=1}^{n_{\text{opt}}}$ distributed according to the initial wave function squared, $|\Psi_{\text{T}}(\boldsymbol{R}; \boldsymbol{\alpha}_0)|^2$ is generated, and then a target function is minimised by varying $\boldsymbol{\alpha}$ at fixed $\{\boldsymbol{R}_i\}$.

The variational energy estimate for this trial wave function may be written

$$E_{\text{VMC}} = \frac{\langle \Psi_{\text{T}} | \hat{H} | \Psi_{\text{T}} \rangle}{\langle \Psi_{\text{T}} | \Psi_{\text{T}} \rangle}, \tag{4.9}$$

which may be used as a target function, as presented in section 3.2.

Another popular target function is the "variance of the VMC energy,"[252,292]

$$\sigma_{\text{VMC}}^2 = \frac{\langle \Psi_{\text{T}} | (\hat{H} - E_{\text{VMC}})^2 | \Psi_{\text{T}} \rangle}{\langle \Psi_{\text{T}} | \Psi_{\text{T}} \rangle}, \tag{4.10}$$

which reaches its minimum of zero when the trial wave function is an eigenstate of the Hamiltonian. In practice, minimising σ_{VMC}^2 yields variational energies, but is affected by large fluctuations, as shown in this chapter.

In continuum QMC methods, modifications have been devised to circumvent this problem, such as weight limiting, unreweighted variance minimization, or the minimization of other measures of spread such as the median absolute deviation from the median energy.[253]

The computational cost of optimizing Jastrow factors within VMC scales as a small power of system size, typically estimated to be $\mathcal{O}(N^3)$.

4.4.1. Variance of the Reference Energy Minimisation

In the context of TC, the reference energy

$$E_{\text{ref}} = \langle D_{\text{HF}} | \mathrm{e}^{-J} \hat{H} \mathrm{e}^{J} | D_{\text{HF}} \rangle \tag{4.11}$$

is of particular significance since it represents the starting point of a TC-FCIQMC calculation (i.e. the walker distributions at imaginary time $\tau = 0$ has this energy). It is also the zeroth-order contribution to the TC-CC energy.

We refer to its associated variance,

$$\sigma_{\text{ref}}^2 = \langle D_{\text{HF}} | \mathrm{e}^{-J} (\hat{H}^{\dagger} - E_{\text{ref}})(\hat{H} - E_{\text{ref}}) \mathrm{e}^{J} | D_{\text{HF}} \rangle \tag{4.12}$$

as the "variance of the reference energy," which is easily evaluated for a finite VMC sample of size n_{opt} as the sample variance of the Slater-Jastrow energy over the HF distribution,

$$S_{\text{ref}}^2 = \frac{1}{n_{\text{opt}} - 1} \sum_{n=1}^{n_{\text{opt}}} \left| \frac{\hat{H}(\boldsymbol{R}_n) \Psi_{\text{SJ}}(\boldsymbol{R}_n)}{\Psi_{\text{SJ}}(\boldsymbol{R}_n)} - \bar{E}_{\text{ref}} \right|^2, \tag{4.13}$$

which tends to σ_{ref}^2 as $n_{\text{opt}} \to \infty$, where $\Psi_{\text{SJ}} := \mathrm{e}^J D_{\text{HF}}$ is the Slater-Jastrow wave function, $\{\boldsymbol{R}_n\}_{n=1}^{n_{\text{opt}}}$ are electronic configurations distributed according to D_{HF}^2, and the VMC estimate of the reference energy is

$$\bar{E}_{\text{ref}} = \frac{1}{n_{\text{opt}}} \sum_{n=1}^{n_{\text{opt}}} \frac{\hat{H}(\boldsymbol{R}_n)\Psi_{\text{SJ}}(\boldsymbol{R}_n)}{\Psi_{\text{SJ}}(\boldsymbol{R}_n)}. \tag{4.14}$$

The variance of the reference energy has been used as a target function for optimising Jastrow factors for the TC method before, albeit in different theoretical frameworks.[223,293]

To understand the physical significance of the variance of the reference energy, note that equation 4.12 may be rewritten as

$$\sigma_{\text{ref}}^2 = \sum_{I \neq \text{HF}} |\langle D_I | \hat{H}_{\text{TC}} | D_{\text{HF}} \rangle|^2, \tag{4.15}$$

where I runs over a complete basis set.[*]

As evident by equation 4.15, minimising σ_{ref}^2 essentially amounts to minimising the coupling of the reference determinant with the remainder of the space, which in the context of FCIQMC translates to a reduced spawning rate from the reference determinant to its connected excited-state determinants, thereby increasing the amplitude of the reference determinant in the resulting CI vector.

Note also that if the Slater-Jastrow wave function were an exact eigenstate of $\hat{H}$, then a TC-FCIQMC simulation starting from the HF determinant would immediately converge to a strictly single-determinant solution. Although this ideal scenario cannot be achieved in practice, it nevertheless illustrates the potential benefits of obtaining a relatively single-reference CI solution by minimising this target function. We expect that this increased single-reference character will also benefit other approaches, particularly those based on single references, such as TC-CC.

We therefore investigate the performance of minimising the variance of the reference as an alternative to minimising the variance of the VMC energy. In figure 4.2, we compare the VMC energy and variance obtained by variance minimisation methods along with energy-minimised[294–296] results for reference, for the systems considerd in this chapter, using $n_{\text{opt}} = 10^5$ VMC configurations.

Minimizing the variance of the VMC energy produces lower average values of σ_{VMC}^2, as one would expect, but also erratic VMC energies with very large standard deviations (up to ~ 50 mHa in our tests). Minimizing the variance of the reference energy, on the other hand, produces values of σ_{VMC}^2 which are only slightly higher on average than those obtained from minimizing the variance of the VMC energy (1–5% in our tests), while producing more stable VMC energies with much smaller standard deviations (up to

---

[*]Here we have a complete basis set because we are optimising with continuum Monte Carlo. If we were to optimise this by directly calculating the matrix elements in equation 4.15, then the sum must be truncated. This is the subject of ongoing work.
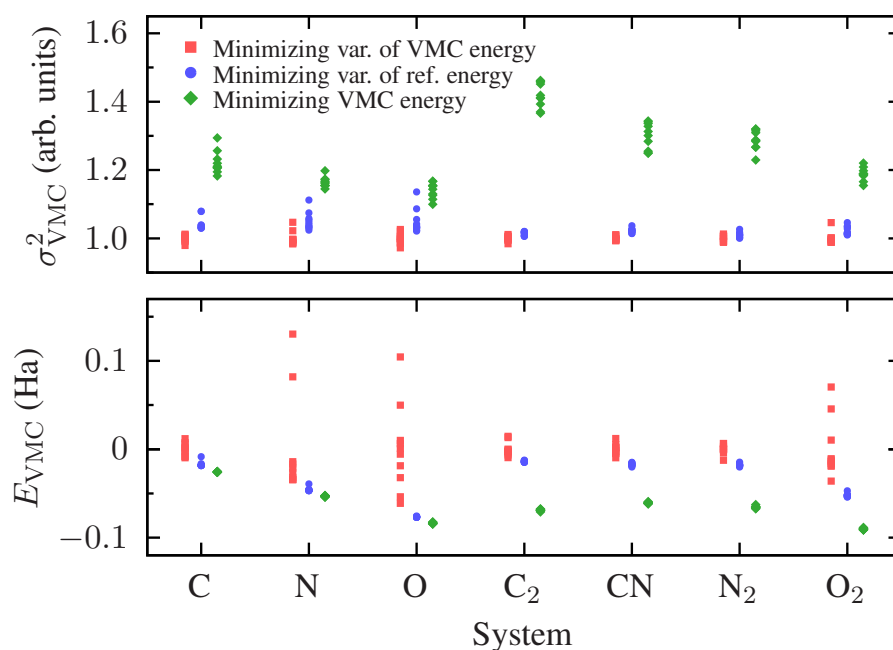
Figure 4.2. Variance of the VMC energy (top) and VMC energy (bottom) of the systems considered in this chapter using the cc-pVTZ basis and Jastrow factors obtained by minimising the variance of the VMC energy (red squares), the variance of the reference energy (blue circles), or the VMC energy (green diamonds) in each of ten independent optimisation runs with $n_{\mathrm{opt}} = 10^5$ VMC configurations. To ease comparisons, variances have been rescaled and energies shifted by their average values from minimising the variance of the VMC energy (i.e. the red squares average to a variance of 1 and an energy of 0 in the plot). The subpar ability of VMC energy variance minimisation to yield consistent VMC energies is evident in the bottom panel, suggesting to use the variance of the reference.

~ 3 mHa in our tests). We therefore do not use "regular" variance minimization since it introduces large stochastic noise, making it unsuitable for optimizing Jastrow factors, and from this point on we use the term "variance minimization" to refer to the minimization of the variance of the reference energy.

4.4.2. Choosing an Appropriate Sample Size

While expectation values relevant for most continuum QMC calculations converge using relatively few VMC configurations, it has also been known[297] that in order to converge other quantities, far more configurations are needed. That is, n_{opt} is larger. In the spirit of MC, we may estimate the convergence of the expectation value for some quantity by performing multiple optimisation runs with different random number seeds but otherwise the same inputs. This would give us a standard deviation.

The value we want to converge in this case is not the VMC energy, nor the reference energy, but instead the TC-FCI energy. In practice, we use the uncertainty of the VMC estimate of the reference energy $\bar{E}_{ref}$ as a proxy for the standard deviation of the TC-FCIQMC energy. This is justified because:

 – The standard deviation of the TC-FCIQMC energy is not larger than the standard deviation of the reference energy, as illustrated in figure 4.3. It is usually significantly smaller thanks to the ability of TC-FCIQMC to conpensate for the presence of a bias in E_{ref} via the correlation energy.

 – The standard deviation of the reference energy is not larger than the statistical uncertainty of the VMC estimate of the reference energy obtained with n_{opt} configurations. It is usually significantly smaller due to the use of variance reduction techniques in QMC.

For the atoms and molecules considered in this chapter, we use $n_{opt} = 2 \times 10^7$ to yield TC-FCIQMC energies with standard deviations of less than 0.1 mHa.

4.4.3. Energy Minimisation

The obvious alternative to variance minimisation is minimising the VMC energy,[294–296] which, as already demonstrated in figure 4.2, results in lower VMC energies but higher VMC variances. Energy minimisation yields wave functions which minimise the statistical fluctuations of the local energy in DMC calculations[298] and is the typical choice for continuum QMC. However, for our purposes, it is unclear whether the resulting wave functions provide a better description of the system than those produced by variance minimisation.

In figure 4.4, we compare the convergence with basis-set size of TC-FCIQMC total energies of the C, N, and O atoms using energy- and variance-minimised Jastrow factors.
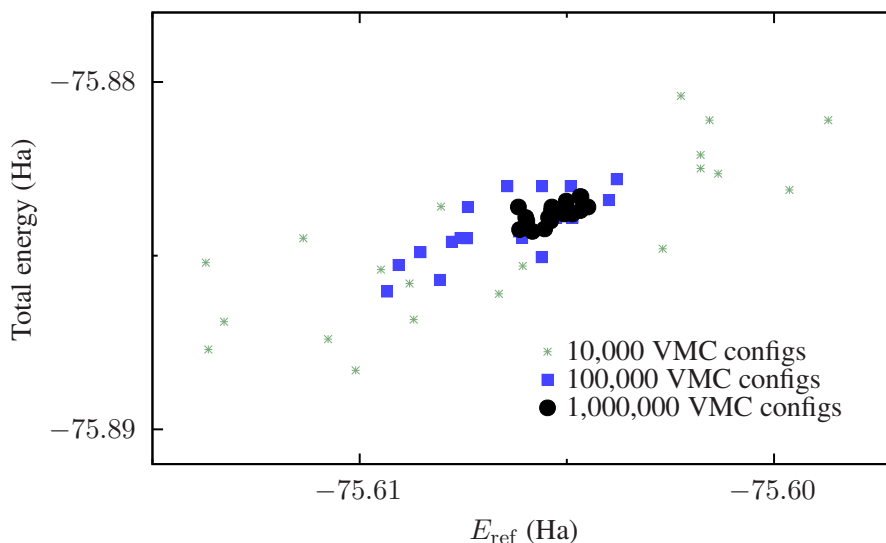
Figure 4.3. TC-FCIQMC energy of the C_2 molecule using 10^6 walkers with the cc-pVDZ basis as a function of the reference energy for multiple independent Jastrow factor parameter sets obtained by variance minimisation using three different VMC sample sizes. The horizonal spread is about 1.8 times larger than the vertical spread, in line with the expectation that the standard deviation of the TC-FCIQMC energy is smaller than that of the reference energy.

Variance minimisation appears to produce wave functions which converge quickly and largely variationally to the basis set limit, while energy-minimised wave functions tend to yield non-variational TC-FCIQMC energies which converge more slowly to the basis set limit.

In figure 4.5 we plot the atomisation energies of the dimers as a function of reciprocal basis-set size, again demonstrating variance-minimised Jastrow factors exhibit favourable convergence properties.

Considering this evidence, we choose to use variance minimisation to optimise the Jastrow factors for subsequent post-HF calculations.

4.5. Grid Sizes

How matrix elements are evaluated is presented in Appendix A. As mentioned there, we use Treutler-Ahlrichs integration grids,[299,300] which are atom-centred grids constructed as the combination of a radial grid running up to the Bragg radius and a Lebedev angular grid. These are obtained using `pyscf`, which provides an integer parameter l_{grid} to control the grid density. Here we test grid errors by evaluating TC-FCIQMC energies at $l_{grid} = 1$–5 and defining the integration error as the difference of each of these results with the value obtained by linear extrapolation to the $1/n_{grid} \to 0$ limit.
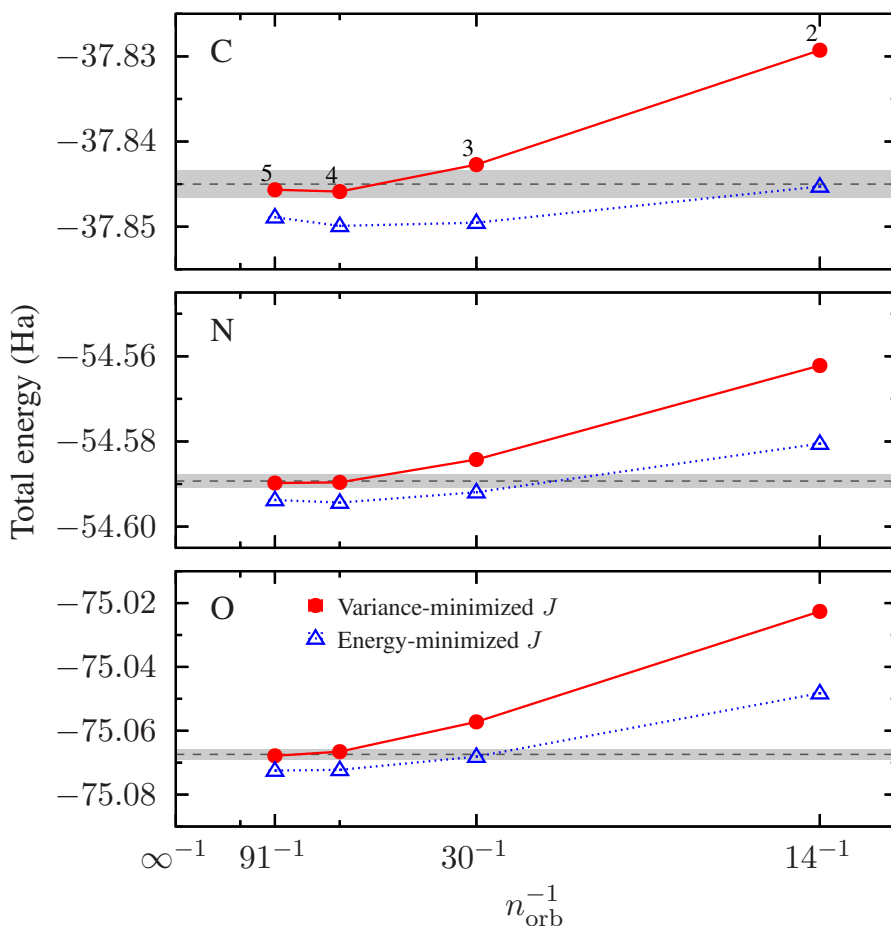
Figure 4.4. Total energy of the C, N, and O atoms as a function of the reciprocal number of molecular orbitals in the cc-pVXZ basis set family. The non-variational behaviour of up to about 5 mHa is evident for the energy-minimised Jastrow factors, for which convergence to the exact energy as a function of basis-set size is slow. The shaded areas represent ± 1 kcal/mol (so-called "chemical accuracy") around the exact non-relativistic total energy from reference 286. Points in the top panel are annotated with the basis set cardinal number X.
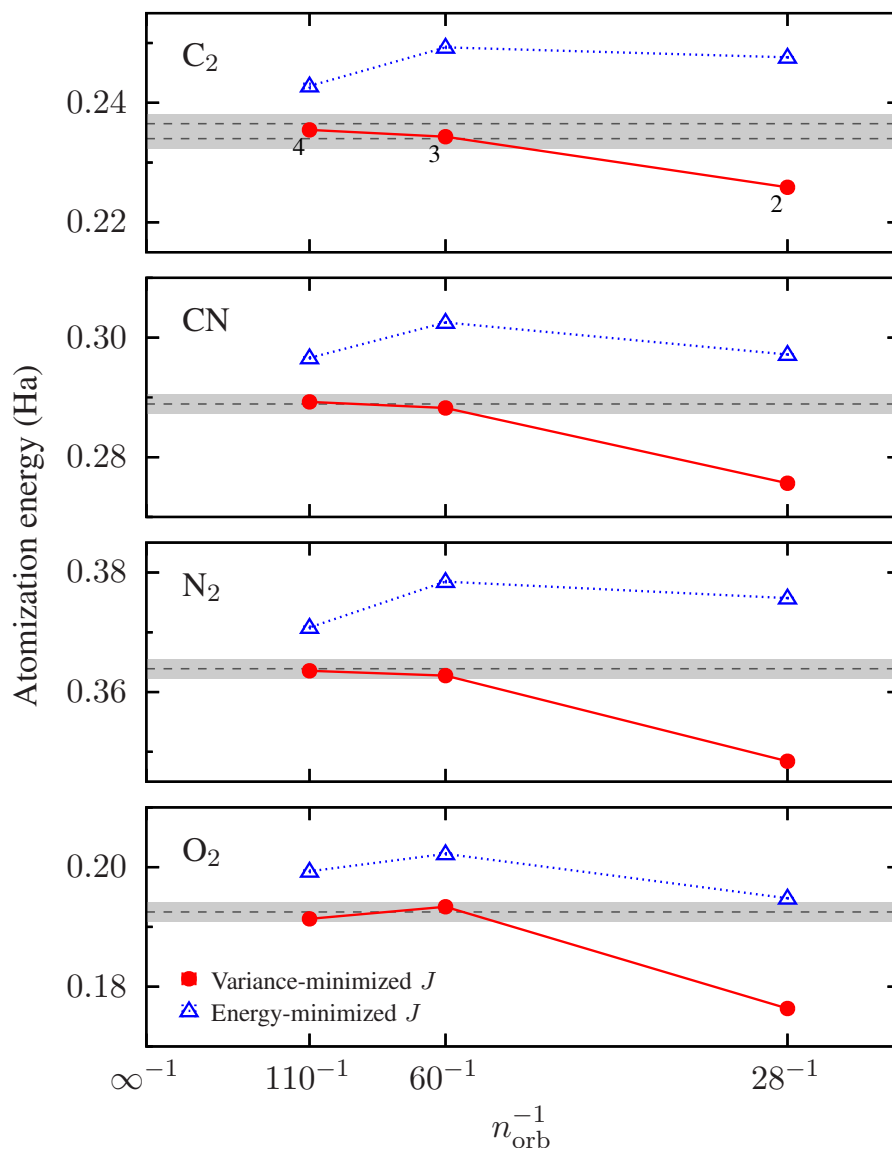
Figure 4.5. Atomisation energy of the C_2, CN, N_2, and O_2 molecules as a function of the reciprocal of the number of molecular orbitals using the cc-pVXZ family of basis sets and Jastrow factors obtained by variance and energy minimisation. The shaded areas represent ± 1 kcal/mol around the theoretical estimate of the exact non-relativistic atomisation energies from references [285, 286]

.

In figure 4.6 we plot the absolute integration error in the total energy of the atoms as a function of $1/n_{\text{grid}}$ for cc-pVDZ, cc-pVTZ and cc-pVQZ basis sets. We find that the basis set size has little to no effect on the integration error.
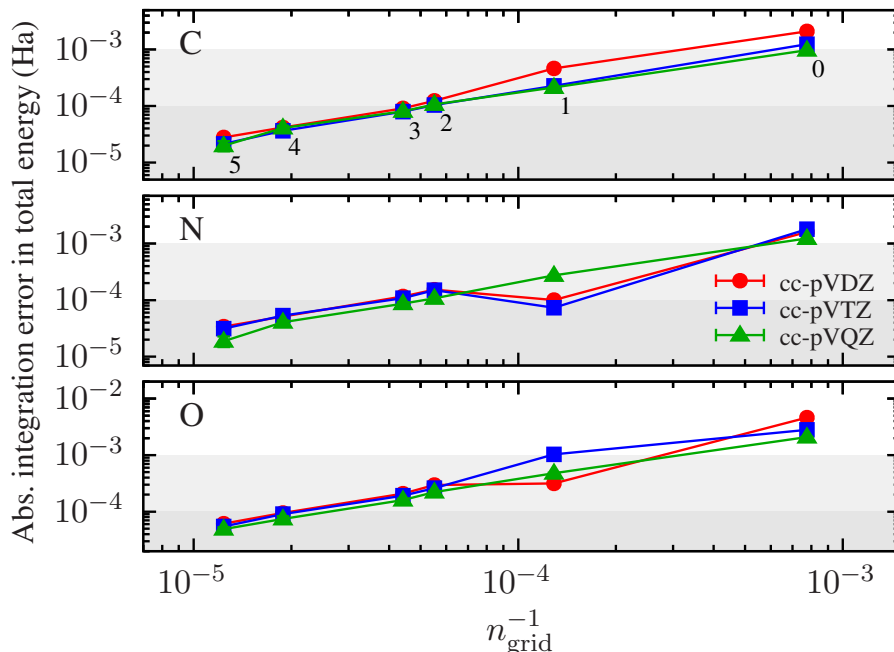
Figure 4.6. Absolute integration errors in the total energy for the C, N, and O atoms as a function of the reciprocal of grid size using basis sets in the cc-pVXZ family. The gray areas correspond to integration errors of less than 1 and 0.1 mHa. Points in the top panel are annotated with the value of `pyscf`'s grid density parameter.

In figure 4.7 we plot the absolute integration error in the total energies of the molecules, the atoms that conform them, and atomisation energies using the cc-pVDZ basis. We find integration-error cancellation in energy differences, with atomisation energy for all four molecules reaching 0.1 mHa for $l_{\text{grid}} = 2$. This represents a 60-point radial grid combined with a 302-point angular grid, for a total of 18120 grid points per atom. We use this throughout this dissertation.

4.6. Compactification of the CI Vector

The more compact the CI wave function is, the easier it is for FCIQMC to sample the wave function accurately and the smaller the initiator error (see section 3.4.5)[282] becomes. The TC method has already been shown to make CI wave functions more compact for the two-dimensional Hubbard model.[301] Let $\{c_I\}$ be the L^2-normalized coefficients of the CI wave function such that $\sum_I c_I^2 = 1$. The quantity

$$\xi = \frac{c_{\text{HF}}^{(\text{TC})} - c_{\text{HF}}^{(\text{non-TC})}}{1 - c_{\text{HF}}^{(\text{non-TC})}} \tag{4.16}$$
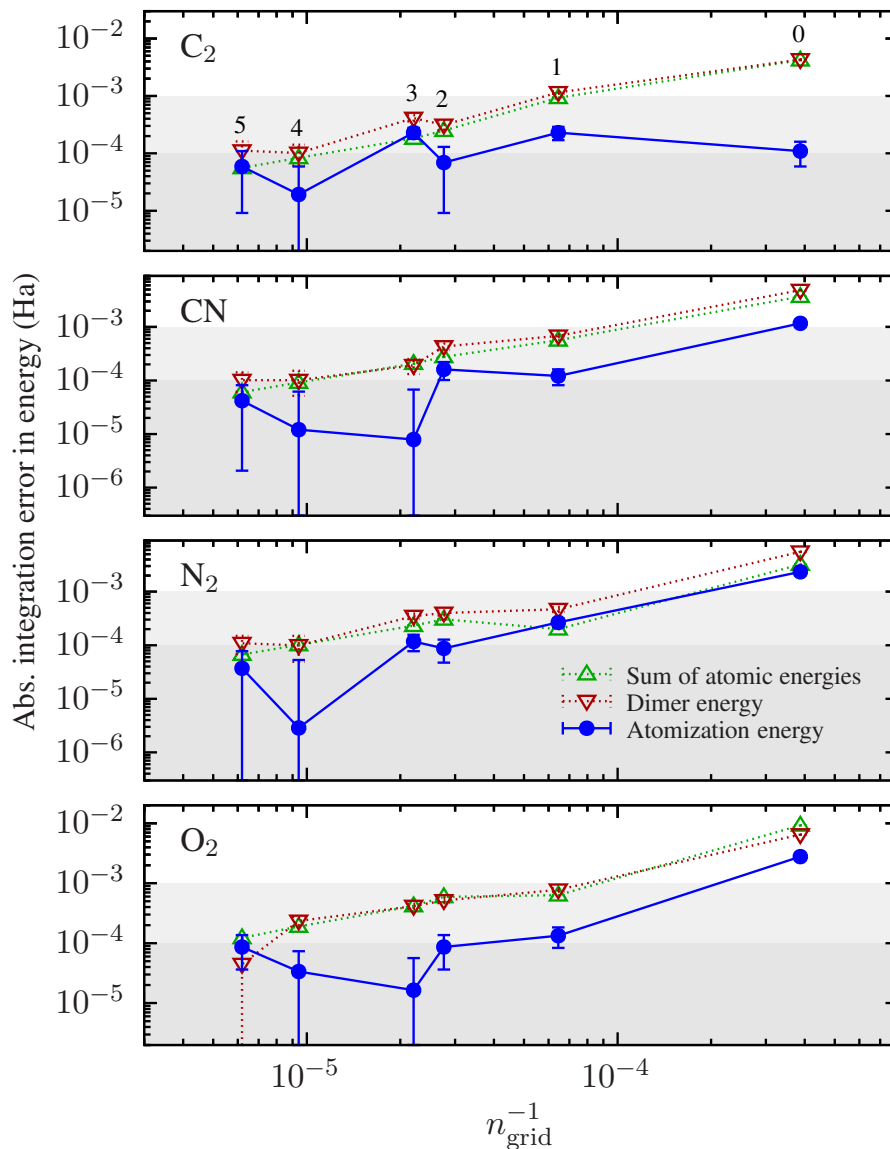
Figure 4.7. Absolute integration error in the energy as a function of the reciprocal number of grid points used for the evaluation of TC integrals. The shaded areas correspond to integration errors of less than 1 and 0.1 mHa. Points in the top panel are annotated with the value of `pyscf`'s grid density parameter. These results demonstrate that $l_{grid} = 2$ is sufficient to achieve sub-mHa accuracy in total energies and sub-0.1-mHa accuracy in relative energies.

is then a measure of the enhancement in the compactness of the wave function, going from 0 for no enhancement to 1 if the TC wave function becomes exactly single-determinantal.

From reference 301, TC yields a maximum of $\xi = 0.64$ for the 18-site 2D Hubbard model. We find that the values of ξ for atomic and molecular systems are not dissimilar from this, as shown in table 4.2.

	C	N	O	C_2	CN	N_2	O_2
cc-pVDZ	0.46	0.63	0.71	0.14	0.23	0.38	0.53
cc-pVTZ	0.45	0.61	0.69	0.15	0.24	0.40	0.55
cc-pVQZ	0.44	0.60	0.69	0.15	0.24	0.41	0.57

Table 4.2. Enhancement of the compactness of the CI wave function, ξ in Eq. 4.16, between our non-TC and TC-FCIQMC calculations.

4.7. Neglecting Three-Body Excitations

As described in section 3.4, sampling in FCIQMC involves spawning walkers from occupied determinants onto connected determinants. The $\hat{L}$ matrix, that is the three-body operator arising from the similarity transformation in TC, results in TC also connecting determinants by three-body excitations. This represents a huge increase in the connectivity of the Hilbert space compared to non-TC-FCIQMC.

These three-body excitations are represented by L_{ijk}^{abc} where each index is distinct. These have been found to typically be small.[192] Moreoever, simultaneous interaction of three electrons is unlikely due to Pauli repulsion (as necessarily at least two of the electrons will be of the same spin). Therefore, we neglect these contributions. The only distinct six-index integrals that we must keep are those occupied in the HF determinant, as these matrix elements are necessary to evaluate the projected energy.

Neglecting pure three-body excitations reduces the amount of storage needed to hold $\hat{L}$ from $\mathcal{O}(M^6)$ to $\mathcal{O}(M^5 + N^3 M^3)$, where M is the number of orbitals and N is the number of electrons. Reduction factors obtains for the molecules studied in this chapter are reported in table 4.3.

	C	N	O	C_2	CN	N_2	O_2
cc-pVDZ	1.23	1.17	1.17	1.87	1.78	1.78	1.58
cc-pVTZ	2.04	2.02	1.93	3.72	3.66	3.66	3.46
cc-pVQZ	3.31	3.44	3.13	6.60	6.54	6.57	6.41

Table 4.3. $\hat{L}$ matrix storage reduction factor from neglecting pure three-body excitations, computed as the number of non-zero matrix elements in a the full $\hat{L}$ matrix divided by the number of non-zero matrix elements with repeated indices or three or more indices corresponding to orbitals occupied in the HF determinant.

Since two-body excitations are typically more expensive to attempt in practice compared to triple excitations, neglecting the latter actually increases the cost per step of the calculation. However, neglecting pure three-body excitations allows the TC-FCIQMC time step to be larger, thereby resulting in reduced serial correlation in the statistics, which enables reaching the target accuracy in fewer time steps, so one can expect a net cost reduction thanks to this approximation. Walltime reduction factors for the molecules studied in this chapter are reported in table 4.4.

	C	N	O	C_2	CN	N_2	O_2
cc-pVDZ	0.9	1.0	1.1	1.7	1.2	1.5	1.6
cc-pVTZ	1.0	1.0	0.8	2.4	1.0	1.8	2.0
cc-pVQZ	1.5	1.5	1.0	3.1	0.9	1.9	2.3

Table 4.4. Reduction factor in the walltime required to advance one unit of imaginary time at fixed population from neglecting pure three-body excitations in the TC-FCIQMC calculation.

The effect of neglecting these pure three-body excitations on atomisation energy is also small, as shown in table 4.5. We find that this approximation results in errors of the order of ~ 0.3 mHa at the cc-pVTZ level, which is a relatively small bias considering the substantial storage and cost benefits of the approximation.

	C_2	CN	N_2	O_2
cc-pVDZ	$-0.62(2)$	$-0.46(0)$	$-0.56(2)$	$-0.55(2)$
cc-pVTZ	$-0.36(5)$	$-0.30(2)$	$-0.32(5)$	$-0.20(3)$
cc-pVQZ	$-0.45(6)$	$-0.21(2)$	$-0.32(7)$	$-0.27(5)$

Table 4.5. Error in the atomisation energy of the molecules considered in this chapter incurred by neglecting pure three-body excitations from the FCIQMC dynamics, in mHa.

4.8. Atomisation Energies

In this section, we compare results of TC-FCIQMC with individually variance-optimised Jastrow factors as a function of basis-set size with the corresponding non-TC results and with benchmark CBS values from references [285–287].

Table 4.6 shows a list of total energies obtained for each system and basis set. We find the TC total energies to be remarkably accurate already at the cc-pVQZ basis set level, differing by less than 2 mHa per atom from benchmark CBS values, while the non-TC total energies still miss the benchmarks by 25–30 mHa per atom with the cc-pV5Z basis set, and 20 mHa per atom at the cc-pV6Z level. The TC total energies exhibit slightly non-variational convergence, with the atomic energies reaching values 0.5 mHa below the benchmark before increasing again towards it for larger basis-set sizes. While

		C	N	O
non-TC	cc-pVDZ	−37.7619	−54.4801	−74.9117
	cc-pVTZ	−37.7900	−54.5252	−74.9853
	cc-pVQZ	−37.8126	−54.5535	−75.0236
	cc-pV5Z	−37.8199	−54.5627	−75.0369
	cc-pV6Z	−37.8263	−54.5697	−75.0447
TC	cc-pVDZ	−37.8293	−54.5622	−75.0226
	cc-pVTZ	−37.8427	−54.5842	−75.0572
	cc-pVQZ	−37.8459	−54.5896	−75.0665
	cc-pV5Z	−37.8457	−54.5898	−75.0678
Ref. 286		−37.8450	−54.5893	−75.0674
Ref. 287		−37.8450	−54.5893	−75.0674

		C_2	CN	N_2	O_2
non-TC	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
TC	cc-pVDZ	−75.8844	−92.6671	−109.4727	−150.2216
	cc-pVTZ	−75.9197	−92.7152	−109.5312	−150.3078
	cc-pVQZ	−75.9272	−92.7247	−109.5428	−150.3244
Ref. 285		−75.9240	−92.7232	−109.5425	−150.3273
Ref. 286		−75.9265		−109.5427	−150.3274
Ref. 287			−92.7229	−109.5425	−150.3275

Table 4.6. Total energies in Ha obtained for atoms (top) and molecules (bottom) considered in this work, along with benchmark non-relativistic results. Statistical uncertainties from Monte Carlo sampling are smaller than 0.0001 Ha.

non-variationality is undesirable in a method, the amount by which the TC results dip below the CBS limit is sufficiently small for this not to be an issue in practice.

From a chemical perspective, relative energies are more important than total energies. The atomisation energies of the C_2, CN, N_2, and O_2 molecules obtained from the total energies in table 4.6 are given in table 4.7 and plotted in figure 4.8.

		C_2	CN	N_2	O_2
non-TC	cc-pVDZ	208.2	255.0	320.7	168.0
	cc-pVTZ	229.4	280.1	351.0	184.9
	cc-pVQZ	232.6	285.6	358.4	189.0
	cc-pV5Z	235.5	289.2	362.6	191.7
TC	cc-pVDZ	225.9	275.6	348.4	176.4
	cc-pVTZ	234.3	288.2	362.7	193.4
	cc-pVQZ	235.5	289.3	363.6	191.4
Ref. 285		234.0	288.9	363.9	192.5
Ref. 286		236.5		364.1	192.6
Ref. 287			288.6	363.9	192.7

Table 4.7. Atomisation energies in mHa obtained for the molecules considered in this chapter, along with benchmark non-relativistic results. Statistical uncertainties arising from Monte Carlo sampling are smaller than 0.1 mHa in all cases.

As with the total energies, we find that the TC atomisation energies converge much faster with respect to basis-set size compared to the non-TC calculations. In fact, the TC results are already chemically accurate at the cc-pVTZ basis-set level, matching accuracy of cc-pV5Z non-TC results. The application of the TC method to quantum chemical methods in general could be presumed to be problematic because any theoretical guarantee of cancellation of errors in energy differences disappears with the introduction of separately-optimized Jastrow factors for each system. However, the fact that in our results the relative energy converges at smaller basis-set sizes than the total energy implies that substantial error cancellation is at play in practice.

4.9. Conclusion and Outlook

In this chapter, we presented a new method to optimise flexible Jastrow factors of a form common in continuum QMC methods for applications to TC. We tested these within the TC-FCIQMC framework. Minimising the variance of the reference energy is shown to be especially good for the TC method, since it maximises the single-reference character of the CI wave function.

These results show that this workflow can provide remarkably accurate total and relative energies, with relative energies at cc-pVTZ rivalling non-TC results at cc-pV5Z. These results also provide better energy estimates compared to other promising methods, such as neural-network based trial wave functions.[302,303]
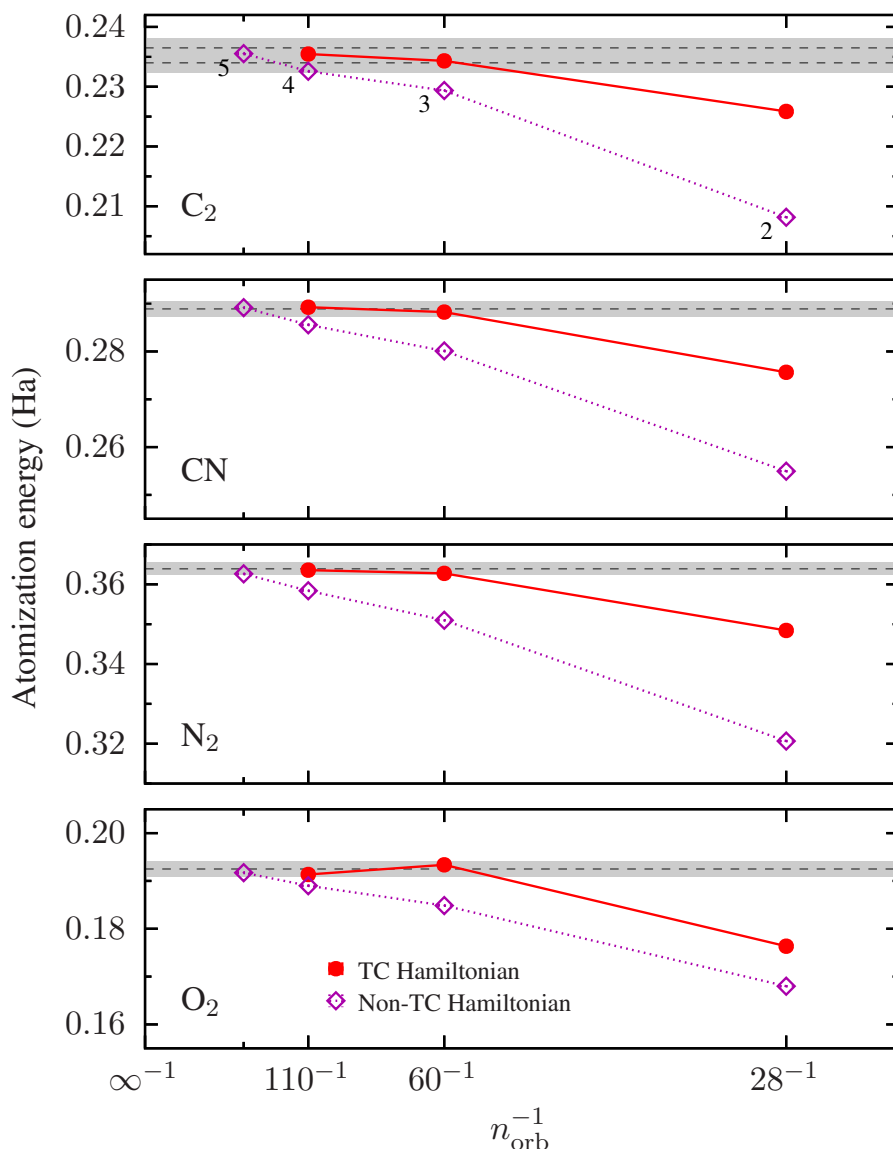
Figure 4.8. Atomisation energy of the C_2, CN, N_2, and O_2 molecules obtained with FCIQMC and TC-FCIQMC as a function of the reciprocal of the number of molecular orbitals using the cc-pVXZ family of basis sets. Points in the top panel are annotated with the basis set cardinal number X. The shaded areas represent ± 1 kcal/mol around the theoretical estimate of the non-relativistic atomisation energy of reference 285; the distinct estimate of reference 286 is also shown for C_2.

Work pursued on this topic have focused on technical advancements and efficient approximations for dealing with the three-body integrals.

4.9.1. The xTC Approximation

One particularly significant refinement of this method is the so-called transcorrelation via exclusion of explicit three-body components (xTC) approximation,[190] wherein upon

neglecting explicit three-body contributions, the remaining three-body terms are "folded" into the remaining lower-order terms.

The Hamiltonian is normal-ordered with respect to Φ (where $\Psi = \mathrm{e}^J \Phi$, with Φ being the HF determinant in this chapter)[304,305]

$$H_N = \hat{H}_{\mathrm{TC}} - \langle\Phi|\, \hat{H}_{\mathrm{TC}}\, |\Phi\rangle = F_N + V_N + L_N \tag{4.17}$$

where the one-, two-, and three-body operators are (using Einstein summation)

$$F_N = \left[h_p^q + \left(U_{pr}^{qs} - U_{pr}^{sq} \right)\gamma_s^r - \frac{1}{2}\left(L_{prt}^{qsu} - L_{prt}^{squ} - L_{prt}^{usq} \right)\gamma_{su}^{rt} \right]\tilde{a}_q^p \tag{4.18a}$$

$$V_N = \frac{1}{2}\left[U_{pr}^{qs} - \left(L_{prt}^{qsu} - L_{prt}^{qus} - L_{prt}^{usq} \right)\gamma_u^t \right]\tilde{a}_{qs}^{pr} \tag{4.18b}$$

$$L_N = -\frac{1}{6} L_{prt}^{qsu}\tilde{a}_{qsu}^{prt} \tag{4.18c}$$

and $U = V - K$, $\tilde{a}_{p\dots}^{q\dots}$ are the normal-ordered excitation operators, and $\gamma_{p\dots}^{q\dots} = \langle\Phi|\, a_{p\dots}^{q\dots}\, |\Phi\rangle$ are the density matrices. The higher-order density matrices are also approximated as the antisymmetrised products of the one-body density matrix. That is, $\gamma_{us\dots}^{rt\dots} \approx \mathcal{A}(\gamma_u^r\gamma_s^t\cdots)$, where $\mathcal{A}$ is the antisymmetriser operator. This means that the only density matrix needed in the xTC approximation is the reduced density matrix (RDM).*

Ignoring the three-body terms L_N leads to an improved scaling with respect to basis set size while directly contracting intermediates to the lower-body corrections obviates the necessity to calculate L and results in excellent energy agreement.[190] Once the integrals are calculated, $\hat{H}_{\mathrm{TC}}$ has, like $\hat{H}$, at most two-body interactions and therefore needs storage for four indices (albeit we no longer have Hermiticity). This means that calculations done on these integrals (in principle) are not much more expensive than the conventional non-TC calculations.

---

*Without this approximation, we would also need the 2RDM and 3-electron reduced density matrix (3RDM). One refinement of the approximation would be to include one or both of these instead of approximating with 1RDMs. Also note that this approximation is exact for a single reference determinant.

The Transcorrelated Method for Multireference Problems

This chapter is based in large part on the following upcoming publication:

Haupt, J. P.; López Ríos, P.; Christlmaier, E. M. C.; Bogdanov, N.; Kats, D.; Alavi, A. "The Transcorrelated Method for Strongly Multireference Problems"

Images have been reused from this paper (with permission).

5.1. Introduction

In this chapter, we apply the new framework for the transcorrated method described in chapter 4 to problems of multireference character and find these methods may yield unphysical results. We propose an updated workflow wherein we use conventional post-Hartree-Fock methods as input to Jastrow factor optimisation for TC-FCIQMC. Calculations suggest size-consistent results and rapid basis set convergence compared to conventional methods, with the binding curve of N_2 at aug-cc-pVTZ being within chemical accuracy of experiment.

5.2. Motivation

A popular "stress test" for quantum chemistry methods is the binding curve of N_2. Highly accurate experimental results[307] exist to recreate the curve, allowing for a useful benchmark. At equilibrium, this system is essentially single reference in character, but as the bond is stretched, the system becomes strongly multireference, making it particularly challenging for many methods. As an example, we might consider standard CCSD, which is a single-reference method, compared to FCIQMC which (within stochastic error) is exact.* Figure 5.1 shows the binding curve of N_2 with CCSD compared to FCIQMC, using cc-pVDZ. We see that CCSD is not stable at large bond lengths.

---

*We do not necessarily need an exact method to resolve this issue. Indeed, coupled cluster may be extended to be multireference, in which case it describes the N_2 binding curve very well.[308]
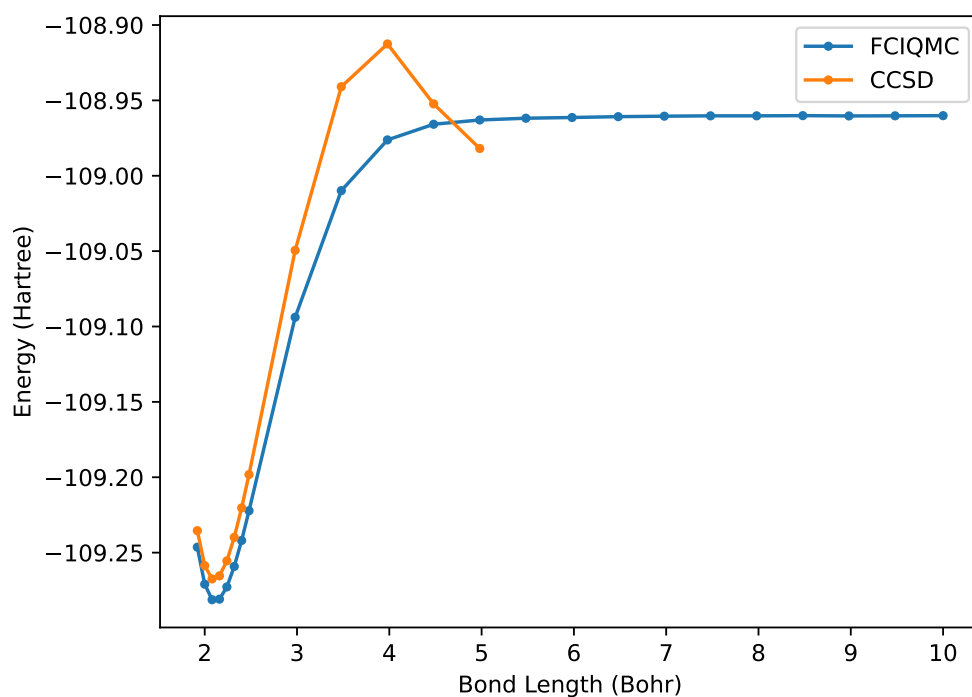
Figure 5.1. The binding curve for N_2 with the cc-pVDZ basis set. CCSD starts to decrease near 4 Bohr, which is unphysical, whereas FCIQMC provides a more accurate curve. This is because CCSD is a single-reference method, whereas FCIQMC is exact (within stochastic error). FCIQMC was done with 30 million walkers and HF-projected energy. CCSD did not converge at bond lengths larger than those shown.

To see how well TC fares against such problems, consider the methodology outlined in chapter 4. We again use the same Jastrow factor as in equation 4.1,

$$J = \sum_{i<j}^{N} v(r_{ij}) + \sum_{i}^{N} \sum_{I}^{N_A} \chi(r_{iI}) + \sum_{i<j}^{N} \sum_{I}^{N_A} f(r_{ij}, r_{iI}, r_{jI}), \tag{5.1}$$

with

$$v(r_{ij}) = t(r_{ij}, L_v) \sum_{k} a_k r_{ij}^k, \tag{5.2}$$

$$\chi(r_{iI}) = t(r_{iI}, L_\chi) \sum_{k} b_k r_{iI}^k, \tag{5.3}$$

$$f(r_{ij}, r_i, r_j) = t(r_{iI}, L_f) t(r_{jI}, L_f) \sum_{k,l,m} c_{klm} r_{ij}^k r_{iI}^l r_{jI}^m, \tag{5.4}$$

and the same cutoff functions $t(r, L) = (1 - r/L)^3 \Theta(r - L)$. We also use the same objective function,

$$\sigma_{\text{ref}}^2 = \sum_{I \neq \text{HF}} |\langle D_I | \hat{H}_{\text{TC}} | D_{\text{HF}} \rangle|^2, \tag{5.5}$$

Using this workflow with the xTC approximation, we calculate points along the binding curve of N_2 with the aug-cc-pVTZ basis set, and find a unphysical "dip", as shown in figure 5.2, similar to what CCSD exhibits in figure 5.1, albeit much more subtle.
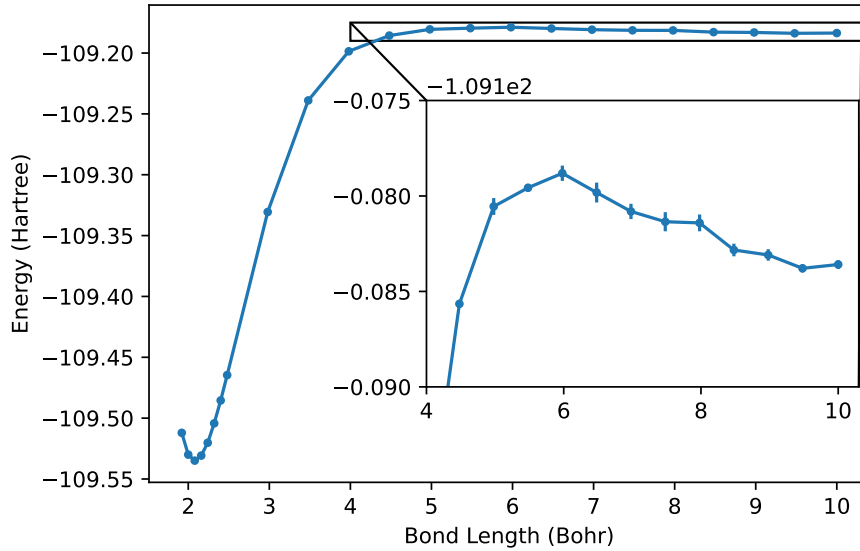
Figure 5.2. The xTC-FCIQMC binding curve for N_2 with the aug-cc-pVTZ basis set. While much smaller than that in figure 5.1, a unphysical dip is still present. This is apparent when zooming in on the curve, as shown in the inset.

This result has been verified for a few points with full TC (that is, with no approximations), which rules out xTC as the issue. Since the post-Hartree-Fock treatment was essentially at the FCI level, this implies that there is something wrong with the calculation

of the Jastrow factors themselves. That is, our Hamiltonian is already "corrupted" before we even start the post-HF calculation.

5.3. Resolving the Problem

Based on the discussion in the previous section, it is likely that the transcorrelated workflow suffers from a single-reference bias. Indeed, the clear culprit is the Slater-Jastrow ansatz

$$\Psi_{\text{SJ}} = e^J \Phi_{\text{HF}}, \tag{5.6}$$

which affects the Jastrow optimisation and in turn the TC Hamiltonian $\hat{H}_{\text{TC}} = e^{-J} \hat{H} e^J$.

We modify equation 5.6 to optimise for a multireference expansion,

$$\Psi = e^J \Phi_0 \tag{5.7}$$

where $|\Phi_0\rangle = \sum_I |D_I\rangle$. In practice, this modification results in two key changes in the workflow:

– The objective function used during the VMC optimisation, equation 5.5, needs to reflect the multireference ansatz. In particular, it will need to be changed to

$$\sigma_{\text{ref}}^2 = \sum_I |\langle D_I| \hat{H}_{\text{TC}} |\Phi_0\rangle - \langle D_I|\Phi_0\rangle \langle \Phi_0| \hat{H}_{\text{TC}} |\Phi_0\rangle|^2, \tag{5.8}$$

which is evaluated in VMC by sampling

$$S_{\text{ref}}^2 = \frac{1}{n_{\text{opt}} - 1} \sum_{n=1}^{n_{\text{opt}}} \left| \frac{\hat{H}(\boldsymbol{R}_n)\Psi(\boldsymbol{R}_n)}{\Psi(\boldsymbol{R}_n)} - \bar{E}_{\text{ref}} \right|^2. \tag{5.9}$$

– The 1RDM, which is used in xTC during integration, must reflect this change as well. In particular, $\gamma_{p\dots}^{q\dots} = \langle \Phi_0| a_{p\dots}^{q\dots} |\Phi_0\rangle$ in the equations in section 4.9.1.

5.3.1. Size Consistency

One possible concern when studying problems like dissociation is that the method be size consistent. It is worth noting that one of the earliest Jastrow factors used for TC[170] as well as in more recent work[175] is given by

$$J = \sum_{i<j} u(\boldsymbol{r}_i, \boldsymbol{r}_j) \tag{5.10}$$

where

$$u(\boldsymbol{r}_i, \boldsymbol{r}_j) = \sum_{l,m,n}^{m+n+o\leq 6} c_{lmn}(\bar{r}_{iM}^m \bar{r}_{jN}^n + \bar{r}_{jM}^m \bar{r}_{iN}^n)\bar{r}_{ij}^l \tag{5.11}$$

and $\bar{r} = r/(1+r)$. However, notice that for $l = 0$ and $n, m > 0$, we can have non-vanishing gradients of u for arbitrary distances between N and M, and hence for systems A and B arbitrarily far apart we do not necessarily have

$$e^{J_{A+B}} |\Phi_{A+B}\rangle = e^{J_0}(e^{J_A} |\Phi_A\rangle)(e^{J_B} |\Phi_B\rangle), \qquad (5.12)$$

as J_A will still act on system B, and vice versa. Here, J_0 is the electron-electron part of the Jastrow factor, J_A the part involving nuclei in system A, and J_B terms involving nuclei in system B.

In contrast to these previous works, our Jastrow ansatz, equation 5.1, first presented by Drummond, Towler and Needs,[309] vanishes when systems A and B are far apart due to the presence of the cutoff functions. Therefore, our Jastrow factor form does not suffer from this problem.

We must also ensure size consistency in our optimisation procedure. Assuming Φ_0 is itself size consistent, it follows that for the given J, for non-interacting systems A and B, the objective function, equation 5.8, $\sigma_{\text{ref}}^2(A + B) = \sigma_{\text{ref}}^2(A) + \sigma_{\text{ref}}^2(B)$, where $\sigma_{\text{ref}}^2(A)$ is the objective function for system A, and similarly for B. Hence, the parameters of the Jastrow factor should converge to the same values when treated as a non-interacting composite system as they would when treated as individual systems. Hence, our updated workflow is size consistent.

5.3.2. Choices for Multireference Ansatzes

We are now faced with the question of which choice of Φ_0 is best. We present here three choices, and discuss their relative merits:

– Using a FCIQMC wavefunction ansatz for Φ_0. This is the most accurate one might get to the true solution, being essentially exact, but it is also computationally prohibitive for large systems. However, it is the most "fool-proof" proof-of-concept choice, can be treated as a blackbox (no need to choose an active space), and we might simply end the calculation early to get the most important components of the CI vector. In this chapter, we use a "snapshot" of the wave function at the end of a non-TC-FCIQMC calculation, and use the associated 1RDM calculated by the "replica trick" presented in section 3.4.6. Since our TC calculation is xTC-FCIQMC, using FCIQMC as the ansatz for Φ_0 might be dubbed "circular" and actually does not worsen the computational scaling of the methodology. Of course, if we choose to use e.g. xTC-DCSD as our TC method, then the computational bottleneck is non-TC-FCIQMC, before we even begin transcorrelation. We will refer to Jastrow factors optimised with this ansatz as "FCIQMC-Jastrows".

– Using a CASSCF wavefunction ansatz for Φ_0. In this method, the orbitals are also modified. It is a compromise compared to the FCIQMC approach, though still quite

costly, is not a blackbox method, and a greater percentage of the wavefunction might be stored in the CI vector (since only a smaller subset of the orbitals are considered). We will refer to Jastrow factors optimised with this ansatz as "CASSCF-Jastrows".

- Using a CASCI wavefunction ansatz for Φ_0. Of the methods presented here, this is the least costly while still potentially capturing much of the static correlation needed and being relatively blackbox.[310] This is probably the most realistic choice for large-scale problems. We will refer to Jastrow factors optimised with this ansatz as "CASCI-Jastrows".

We refer to the Jastrow factors optimised with the restricted HF ansatz described in the previous section as "RHF-Jastrows".*

5.4. Trial Wavefunctions in TC-FCIQMC

Another challenge when studying multireference problems with FCIQMC is noise in the projected energy. Here we generalise the discussion on trial wavefunctions presented in section 3.4.3.

Consider a TC-FCIQMC calculation, where we wish to solve for the eigenvalue of the ground state Φ for the TC Hamiltonian $\hat{H}_{\mathrm{TC}}$. Conventionally, the projected energy can be written

$$E_{\mathrm{proj}} = \frac{\langle \Phi_{\mathrm{trial}} | \hat{H}_{\mathrm{TC}} | \Phi_{\mathrm{FCIQMC}} \rangle}{\langle \Phi_{\mathrm{trial}} | \Phi_{\mathrm{FCIQMC}} \rangle} \tag{5.13}$$

where $|\Phi_{\mathrm{FCIQMC}}\rangle$ is the estimate of the wave function according to the FCIQMC algorithm, and $|\Phi_{\mathrm{trial}}\rangle$ is the trial wave function. If we write $|\Phi_{\mathrm{FCIQMC}}\rangle$ as a sum of the exact wave function $|\Phi\rangle$ plus some error $|\delta\rangle$, we have

$$E_{\mathrm{proj}} = \frac{\langle \Phi_{\mathrm{trial}} | \hat{H}_{\mathrm{TC}} | \Phi + \delta \rangle}{\langle \Phi_{\mathrm{trial}} | \Phi + \delta \rangle} \tag{5.14}$$

$$= \frac{E_0 \langle \Phi_{\mathrm{trial}} | \Phi \rangle + \langle \Phi_{\mathrm{trial}} | \hat{H}_{\mathrm{TC}} | \delta \rangle}{\langle \Phi_{\mathrm{trial}} | \Phi \rangle + \langle \Phi_{\mathrm{trial}} | \delta \rangle} \tag{5.15}$$

where E_0 is the exact ground-state energy. If $|\Phi_{\mathrm{trial}}\rangle$ is the left eigenvector, then $\langle \Phi_{\mathrm{trial}} | \hat{H}_{\mathrm{TC}} | \delta \rangle = (\hat{H}_{\mathrm{TC}}^\dagger | \Phi_{\mathrm{trial}} \rangle)^\dagger | \delta \rangle = E_0 \langle \Phi_{\mathrm{trial}} | \delta \rangle$, so

$$E_{\mathrm{proj}} = \frac{E_0 \langle \Phi_{\mathrm{trial}} | \Phi \rangle + E_0 \langle \Phi_{\mathrm{trial}} | \delta \rangle}{\langle \Phi_{\mathrm{trial}} | \Phi \rangle + \langle \Phi_{\mathrm{trial}} | \delta \rangle} \tag{5.16}$$

$$= E_0 \frac{\langle \Phi_{\mathrm{trial}} | \Phi \rangle + \langle \Phi_{\mathrm{trial}} | \delta \rangle}{\langle \Phi_{\mathrm{trial}} | \Phi \rangle + \langle \Phi_{\mathrm{trial}} | \delta \rangle} \tag{5.17}$$

$$= E_0, \tag{5.18}$$

---

*Unrestricted HF along with spin-dependent Jastrow factors are the subject of a future work.

i.e. if our trial wavefunction is the left eigenvector of $\hat{H}_{\mathrm{TC}}$, then $E_{\mathrm{proj}} = E_0$.

In standard FCIQMC, we take the top N_T determinants at some point of a calculation, form a subspace by constructing a $N_T \times N_T$ trial space H_{trial}, diagonalise it exactly, and use the eigenvector as an approximation to the exact one, to be used as $|\Phi_{\mathrm{trial}}\rangle$ in the calculation of the projected energy. As H is Hermitian, taking the left or the right eigenvector is irrelevant. To get the left eigenvector in an equivalent way would involve doing a FCIQMC calculation on $\hat{H}_{\mathrm{TC}}^{\dagger}$, but this is costly and instead we find taking the left eigenvector of the subspace from $\hat{H}_{\mathrm{TC}}$ to be a good approximation, as the two typically have similar top determinants. Moreover, this method gives the correct energy as long as the trial wavefunction has nonzero overlap with the true eigenvector, so even if our choice is not perfect, it will still give the correct answer, and likely with a smaller error compared to using the Hartree-Fock determinant.

As an example of the effectiveness of this approach, consider a highly dissociated point in the binding curve of the nitrogen molecule. This is shown in figure 5.3, specifically at 10 Bohr with the aug-cc-pVTZ basis set and 10^8 walkers. For this calculation, FCIQMC-Jastrows were used. We can see from this figure that using the trial-projected energy even
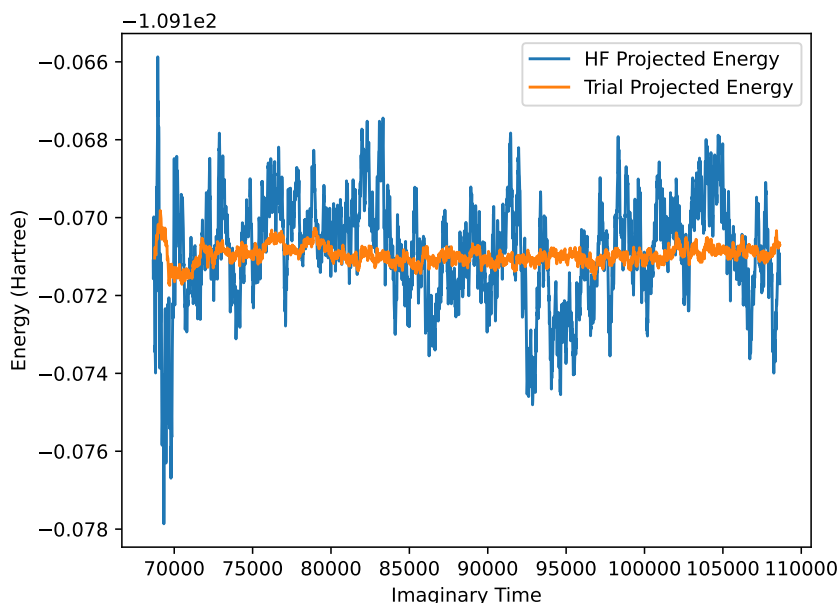
Figure 5.3. The HF-projected and trial-projected energy trajectories in imaginary time for N_2 with a separation of 10 Bohr with the aug-cc-pVTZ basis set and the Jastrow factor optimised for the variance of the FCIQMC energy. The calculation was done with 10^8 FCIQMC walkers. This is a highly dissociated and hence multireference state. The trial-projected energy uses the top 10 determinants but substantially improves the rate of convergence when compared to the HF projected energy.

with just a few determinants allows us to much more easily handle highly multireference problems. Thus, trial wave functions are used throughout the rest of this chapter.

5.5. Results

5.5.1. Computational Details

We calculate the energy of the nitrogen molecule across multiple bond lengths, ranging from 1.92 to 10 Bohr. We use the aug-cc-pVTZ basis set, which contains diffuse functions for long-range correlations while still being a modest size. We calculate the non-TC CI vectors and 1RDMs for each geometry with FCIQMC (using `NECI`),[267] CASSCF (using `Molpro`),[311–313] and CASCI (using `pyscf`).[289] The CI vector is then used in the objective function for optimising the Jastrow factor with VMC using `CASINO`.[253] Using the optimised Jastrow factor and 1RDM, we then calculate the relevant integrals for the xTC Hamiltonian using `pytchint`. Finally, xTC-FCIQMC is performed on these integrals with `NECI`. Each geometry is calculated independently; that is, the Jastrow factor is optimised for each geometry separately. In order to keep memory usage manageable, we cutoff the number of determinants in our CI vector to be 100. Even at the dissociated limit, the number of relatively-highly-weighted determinants is around 20, so 100 should be enough to capture the static correlation for the VMC optimisation.

Next, we also calculate the vertical excited-state energies using this workflow for a few states of the nitrogen molecules with CASSCF- and CASCI-Jastrows. Excited states are also challenging multireference problems. Moreover, we optimise the Jastrow factors in a state-specific manner. That is, for some excited state Φ_{exc}, our ansatz becomes $\Psi_{\mathrm{exc}} = \mathrm{e}^J \Phi_{\mathrm{exc}}$, thereby modifying our workflow slightly to optimise specifically for that state, as well as using its 1RDM for the xTC approximation. Naturally, the xTC-FCIQMC calculation will be targeting this state. For cases where a triplet excited state of the same symmetry is lower in energy than a singlet excited state, the FCIQMC calculation will collapse to the triplet state. To overcome this, we use a spin-penalty term to target the singlet excited state.[314]

For the N_2 binding curve, we compare against the highly accurate experimental curve determined in reference 307, and as a benchmark we compare against F12 calculations, performed in `Molpro`.[311–313] For excited states we compare against extrapolated FCI calculations reported in reference 315.

5.5.2. Binding Curves

As illustrated in figure 5.4, we obtain a qualitatively-correct binding curve for the nitrogen molecule using any of the Jastrow factors, besides the RHF-Jastrow. Shown in that figure is also a zoom-in on the dissociated limit, indicating the RHF-Jastrow curve is the only one exhibiting the pathological "dip" behaviour. The remaining unphysical behaviour amounts to noise, which has a few potential sources:

  – The optimisation of the Jastrow factor is done with VMC, a stochastic algorithm. As described in chapter 4, VMC optimisation for the TC method is complex, and as

of yet has not been done on strongly multireference problems. Since the optimisation is done independently for each point along the binding curve, this may lead to some noise.

– The FCIQMC calculations are also stochastic, and this may also lead to some noise, particularly in the dissociated limit when not using a multideterminantal trial wave function.

– In the case of the FCIQMC-Jastrow, even the non-TC calculation prior to Jastrow optimisation is stochastic. In this case, the CI vector and 1RDM collected from the non-TC-FCIQMC calculation is done so at only a snapshot in imaginary time: right at the end of the calculation. One way to reduce this noise (not explored here) would be to average these values over imaginary time.

Thus, we can conclude that the biggest problem with the TC workflow has been resolved in three ways by introducing a multireference Jastrow optimisation (and 1RDM for the xTC approximation).

We now consider the binding curve at the equilibrium geometry (taken to be 2.08 Bohr radii). Such a zoom-in is shown in figure 5.5. Since N_2 is largely single-reference in this region, we expect all of the curves, including that with the RHF-Jastrow, to approximately be equal there. Relative to the RHF-Jastrow curve,

– The CASCI-Jastrow curve uses the same orbitals and uses only a small subspace of the full CI space, and is indeed strongly single-reference in this region. As a result, this curve is almost identical to the RHF-Jastrow curve in the region around the minimum.

– The CASSCF-Jastrow curve has a noticeable shift of about 2.8 mHa. While not expected, since the TC method is nonvariational we cannot guarantee a more complete wave function description to necessarily lead to a lower energy value. However, notice in the bottom panel of figure 5.4 the CASSCF-Jastrow curve is shifted upwards relative to the CASCI-Jastrow curve by about the same amount, so in relative terms this is actually acceptable.

– The FCIQMC-Jastrow is slightly higher in energy compared to the RHF- and CASCI-Jastrow curves near the minimum, and similarly above the CASCI-Jastrow curve at dissociation. Since it shares the same orbitals as the CASCI-Jastrow curve, and we are primarily concerned with static correlation at this part of the workflow, we expect the FCIQMC-Jastrow curve to have a similar shape, and indeed it does. Possible sources of differences are: stochastic noise and having more flexibility to capture the full wave function (especially relevant at dissociation).
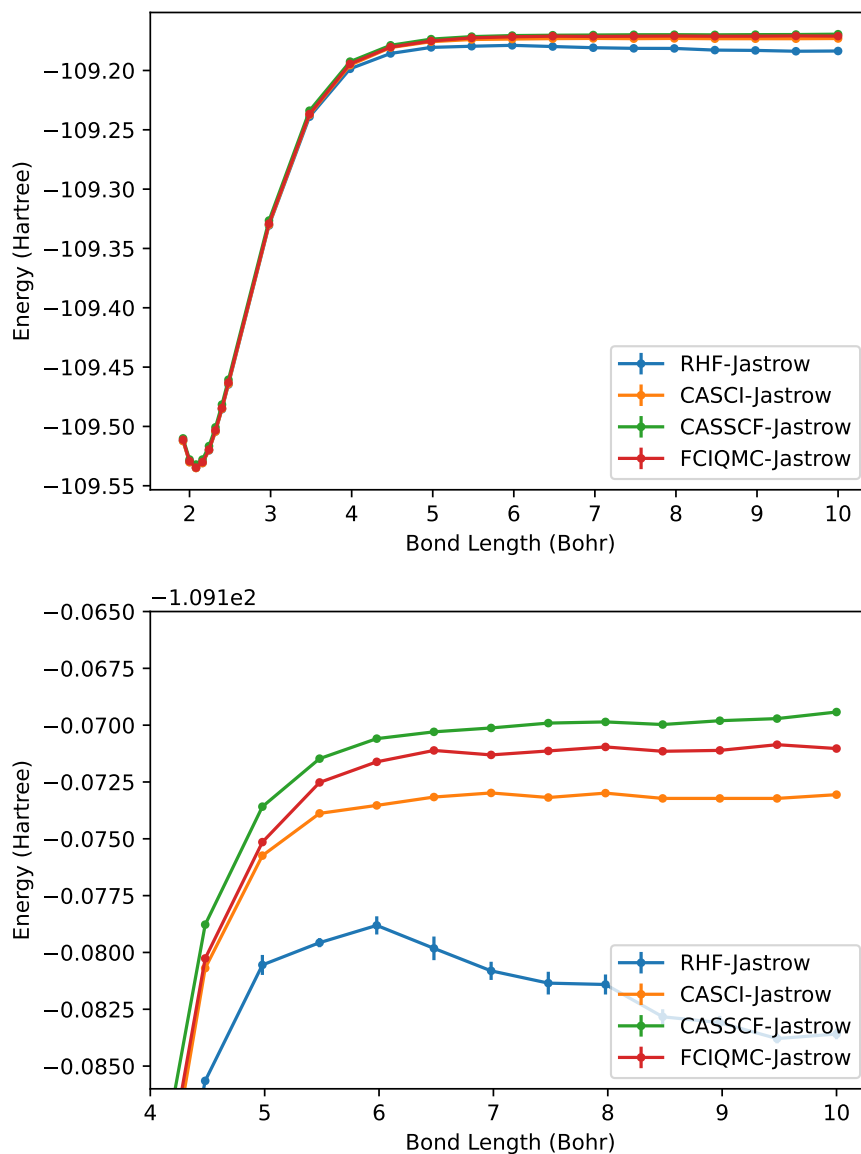
Figure 5.4. xTC-FCIQMC energies for the nitrogen dimer for various points along its binding curve, between 1.92 and 10 Bohr radii. Calculations were performed with the aug-cc-pVTZ basis set. Four choices for Jastrow factors are presented. The forms for the actual Jastrow factor J is the same, but the value for $|\Phi\rangle$ in the ansatz $|\Psi\rangle = \mathrm{e}^J |\Phi\rangle$ is different. The choices are: RHF-Jastrow (blue), CASCI-Jastrow (orange), CASSCF-Jastrow (green) and the FCIQMC-Jastrow (red). The top panel shows the full binding curve, while the bottom panel shows the dissociated limit. Notice that except for the RHF-Jastrow curve, these Jastrow factors result in qualitatively-correct xTC-FCIQMC binding curves. Noise near dissociation is likely due to the VMC optimisation, which was shown in chapter 4 to have an error in the final TC energy of about 0.1 mHa.
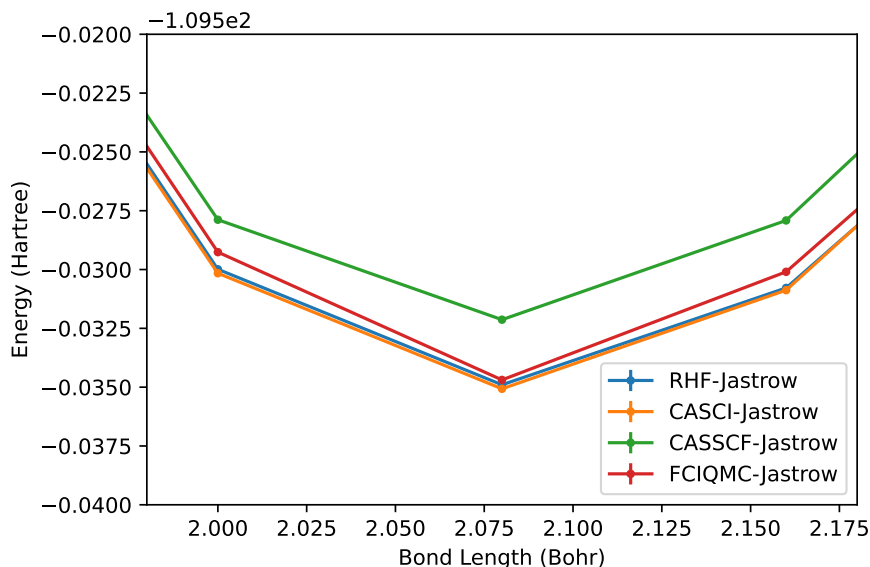
Figure 5.5. The N_2 binding curves at aug-cc-pVTZ for the RHF- (blue), CASCI-
(orange), CASSCF- (green) and FCIQMC-Jastrow (red), zoomed in near equilibrium.
Here, the problem is strongly single-reference and hence we expect all curves to be similar.
However, the CASSCF-Jastrow curve is shifted upwards relative to the CASCI-Jastrow
curve by about 2.8 mHa, but this is compensated for at dissociation.

Naturally, we also wish to determine how accurate our calculations are. This is easy to
verify for the extremes, as we can directly compare against high-accuracy extrapolated
ab initio thermochemistry (HEAT).[285] However, we also want to evaluate the accuracy of
the rest of the binding curve. For this, we test against an accurate fit to experimental
spectroscopic data which includes high-energy dissociated data.[307] Since we are only
interested in relative energies, we normalise all curves to be zero at 10 Bohr radii, and
subtract the experimental values. The result of this is shown in figure 5.6.

We find that the CASCI- and FCIQMC-Jastrow curves are within chemical accuracy
(1.6 mHa) for most of the binding curve and the CASSCF-Jastrow entirely within chemical
accuracy when compared against experiment, all of them outperforming MRCI-D-F12. It
is worth noting, however, that unlike MRCI-D-F12, our multireference Jastrow factors
are not uniformly above experiment. As expected, the RHF-Jastrow performs poorly.

Another worthwhile check is how accurate the atomisation energy is, taking the
extremely dissociated limit (10 Bohr radii) and comparing it against equilibrium (2.08
Bohr radii). This is shown in table 5.1. Relative to experiment, all multideterminantal
Jastrow factors are within chemical accuracy (1.6 mHa).

Finally, as a measure for the size consistency error, we consider the difference between
twice the atom's energy and that of the molecule at a separation of 10 Bohr radii. This
is shown in table 5.2. We find that the CASSCF- and FCIQMC-Jastrow factors are
size consistent to a reasonable degree (similar to MRCI-D-F12) whereas the CASCI-
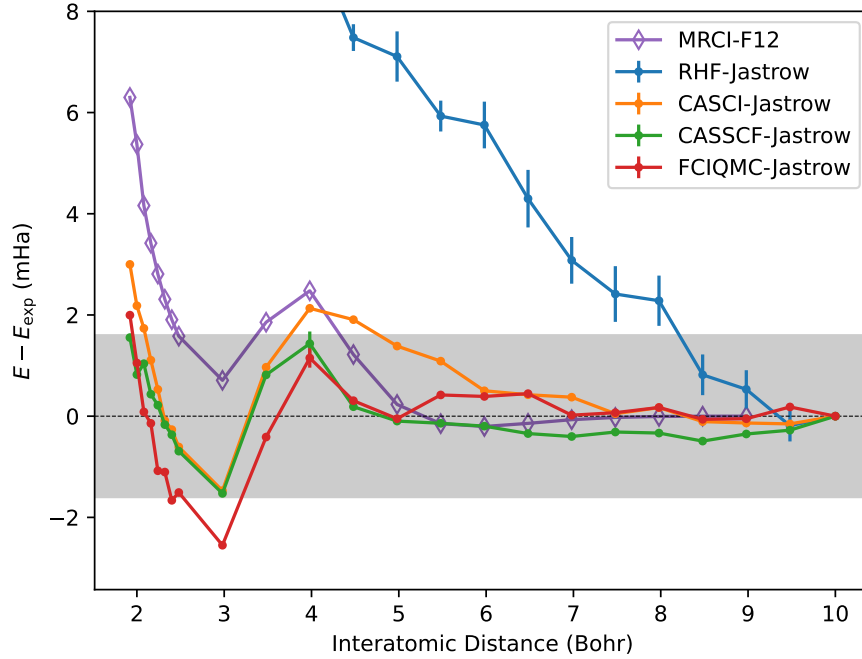
Figure 5.6. Binding curves for each Jastrow factor optimisation strategy relative to the experimental fit from reference 307. All curves are normalised such that the energy at 10 Bohr radii is zero, except for MRCI-D-F12 where energy is set to zero at 8.98 Bohr radii (past this, it did not converge). The shaded region represents 1.6 mHa, so-called "chemical accuracy". We see that most of the curves for the multirefence Jastrow optimisations are within chemical accuracy, outperforming MRCI-D-F12. Note that while the relative values are below experiment for some regions (notably around 3 Bohr), all absolute values are above the HEAT result for N_2 at equilibrium.

Jastrow Factor	Atomisation Energy (mHa)
CASCI	362.0
CASSCF	362.7
FCIQMC	363.6
MRCI-D-F12	359.5
HEAT[285]	363.9
Experiment[307]	363.7

Table 5.1. Atomisation energies using the binding curves for the multideterminantal Jastrow factor optimisation choices. Relative to experiment, all TC calculations are within chemical accuracy (1.6 mHa), with the CASI-Jastrow narrowly outside it relative to HEAT. The RHF-Jastrow is not included in the table because it does not stabilise to a dissociated limit. Note also that these values have an additional error coming from VMC of about 0.3 mHa, according to the study from chapter 4 (0.1 mHa error for the molecule, and for each atom). All multideterminantal Jastrow factors outperform MRCI-D-F12 according to this measure.

Jastrow factor is not as size consistent. Non-TC CASCI was already size inconsistent, and the Jastrow optimisation was likely not able to adequately compensate for the missing dynamical correlation.

Jastrow Factor	$\sum E_{\text{atom}} - E_{\text{molecule}}(r = 10)$ (mHa)
CASCI	2.1
CASSCF	-1.5
FCIQMC	0.5
MRCI-D-F12	-0.9

Table 5.2. Size consistency error for the multideterminantal Jastrow optimisation choices, expressed as the difference between twice the energy of the atom and the energy of the molecule at dissociation. CASSCF- and FCIQMC-Jastrow factors show size consistency similar to MRCI-D-F12, which we use as a benchmark. The CASCI-Jastrow is has a larger error, but this is likely because the non-TC CASCI was already size inconsistent, and the Jastrow optimisation was not able to adequately compensate for the missing dynamical correlation. Including the RHF-Jastrow factor is meaningless because it does not stabilise in the dissociated limit.

5.5.3. Excitation Energies

One major advantage to this updated workflow is the ability to explicitly target specific states. We demonstrate this by calculating select excitation energies of the nitrogen molecule. We use two approaches to generate the Jastrow factors J, both based on the full valence $(10o, 8e)$ active space:

– CASCI$(10o, 8e)$-J, where the CI vector is chosen "minimally"; that is, we take 99% of the total wave function as measured by the squared modulus of the coefficients (and including additional determinants if degenerate). This, for example, results in only the RHF determinant in the ground state, and only two in the excited states, plus possibly some very small additional determinants being included in the CI vector. This approach does no additional optimisation of the orbitals beyond RHF.

– State-averaged (SA)-CASSCF$(10o, 8e)$-J, where all states are included in the orbital optimisation (therefore, all states share the same orbitals). For each state, we use the top 100 determinants of the CI vector as the Φ ansatz.

The CASCI or CASSCF calculations are used to then determine the 1RDM for the state. The CI vector is then used in the VMC optimisation, and the 1RDM for the xTC approximation when calculating integrals. For simplicity, we consider only those states which are the lowest in their symmetry, which allows us to use standard ground-state FCIQMC for the excited states.[*]

---

[*]In principle, we may also target higher-order excited states, but this would require an additional adjoint replica.[316]
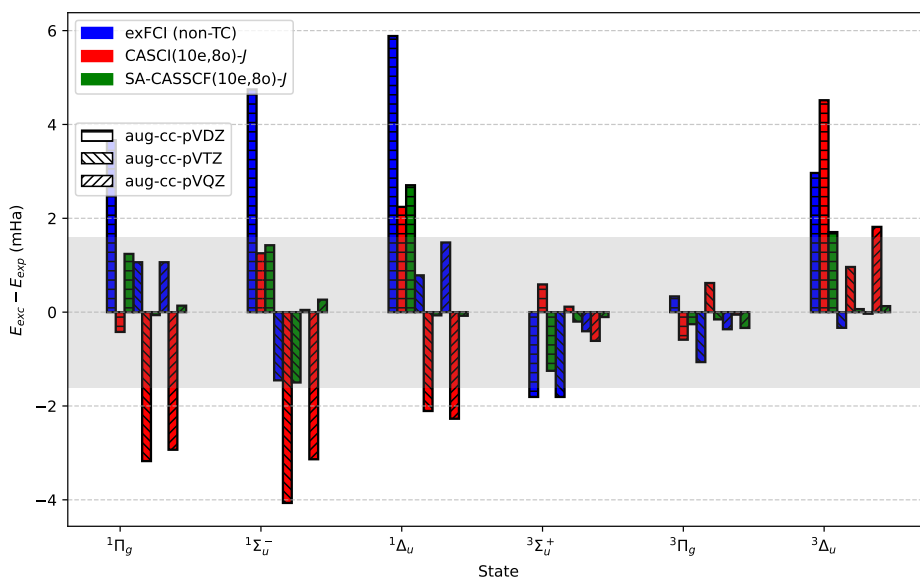
Figure 5.7. Excitation energies E_{exc} for each state compared to experiment E_{exp},[317,318] $E_{\text{exc}} - E_{\text{exp}}$ for the nitrogen dimer. Shown are the results from extrapolated FCI (exFCI), taken from reference 315, xTC-FCIQMC with CASCI-Jastrow factors and xTC-FCIQMC with state-averaged CASSCF-Jastrow factors for the aug-cc-pVDZ (horizontal stripes), aug-cc-pVTZ (backslash hatch pattern) and aug-cc-pVQZ (forward-slash hatch pattern) basis sets. The SA-CASSCF Jastrow factors are shown to largely outperform the other two approaches, being generally within chemical accuracy, even for relatively modest basis sets. The CASCI Jastrow factors perform unfavourably, however. This could be because the orbitals are not optimised for the active space, or it could be because the CI vector is shorter, and the number of determinants for each state is not consistent. Therefore, it is reasonable to assert that the use of the transcorrelated method with optimised orbitals and tailored Jastrow factors allow for highly accurate excitation energies.

Figure 5.7 displays the excitation energies E_{exc} of the N_2. Of the three methods explored (xTC-FCIQMC with SA-CASSCF- or CASCI-Jastrow factors, and non-TC), SA-CASSCF-Jastrow factors give the most consistently accurate excitation energies. The SA-CASSCF-Jastrow factors result in excitation energies typically within chemical accuracy even with aug-cc-pVDZ, and consistently in chemical accuracy for aug-cc-pVTZ and aug-cc-pVQZ. This suggests a SA-CASSCF is a particularly fruitful ansatz for Φ when used in the TC method to calculate excited states.

Dinitrogen

State	Method	aug-cc-pVDZ	aug-cc-pVTZ	aug-cc-pVQZ	Benchmarks
$^1\Pi_g$	non-TC exFCI[a]	345.8	343.2	343.2	344.47[b]
	CASCI-xTC	341.7	339.0	339.2	344.47[c]
	SA-CASSCF-xTC	343.4	342.1	342.3	342.99[d]
$^1\Sigma_u^-$	non-TC exFCI[a]	369.3	363.1	364.6	367.04[b]
	CASCI-xTC	365.8	360.5	361.4	367.04[c]
	SA-CASSCF-xTC	366.0	363.1	364.8	373.33[d]
$^1\Delta_u$	non-TC exFCI[a]	383.3	378.2	378.9	379.99[b]
	CASCI-xTC	379.7	375.3	375.1	379.99[c]
	SA-CASSCF-xTC	380.1	377.3	377.3	389.98[d]
$^3\Sigma_u^+$	non-TC exFCI[a]	283.0	283.0	284.4	286.75[b]
	CASCI-xTC	285.4	284.9	284.2	286.75[c]
	SA-CASSCF-xTC	283.6	284.6	284.7	279.72[d]
$^3\Pi_g$	non-TC exFCI[a]	295.8	294.4	295.1	297.48[b]
	CASCI-xTC	294.9	296.1	295.4	297.48[c]
	SA-CASSCF-xTC	295.2	295.3	295.1	297.85[d]
$^3\Delta_u$	non-TC exFCI[a]	329.3	326.0	326.3	328.56[b]
	CASCI-xTC	330.8	327.3	328.2	328.56[c]
	SA-CASSCF-xTC	328.0	326.4	326.5	330.41[d]

[a] Taken from reference 315.
[b] Experimental value, see references [315, 317, 318].
[c] Experimental value, see references [315, 318, 319].
[d] Theoretical (multireference coupled cluster) value, see references [315, 320].

Table 5.3. Excitation energies for the nitrogen dimer various excited states, in mHa. We use xTC-FCIQMC with CASCI- or SA-CASSCF orbitals and CI vector ansatz for the TC method, and compare it to experiment and non-TC results. We find that while no method consistently beats all others, SA-CASSCF is a particularly effective choice for calculating excited states in the context of TC.

The excitation energies are presented in tabular format in table 5.3. Here we can see again the relative efficacy of SA-CASSCF orbitals and CI vectors for use for the TC method, compared to experimental values.

CAS-Only TC Excitation Energies

Since CASSCF primarily serves to capture the effects of *static* correlation, whereas TC is designed to capture the effects of *dynamical* correlation, it is interesting to see how well the method performs when we perform our calculations only inside the active space. For this, we use the SA-CASSCF$(10e, 8o)$-Jastrow factors from before, but instead of performing all-electron xTC-FCIQMC, we instead only diagonalise the Hamiltonian in the active space. That is, we perform a xTC-CASCI$(10e, 8o)$ calculation, using the orbitals optimised by the non-TC SA-CASSCF calculation. The xTC-CASCI calculations were performed using NECI.[267]
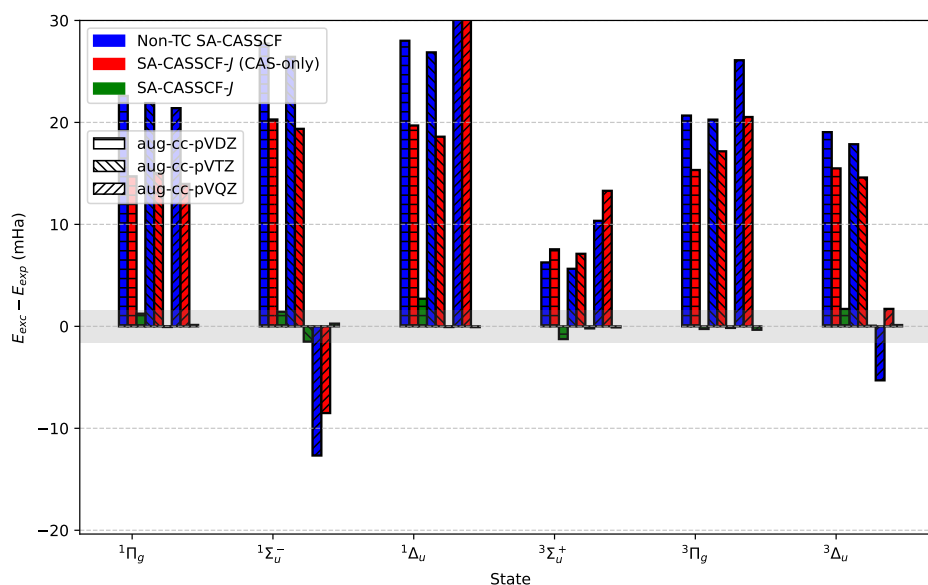
Figure 5.8. Excitation energies relative to experiment for a few states of N_2. For comparison, we show using non-TC SA-CASSCF, as well as xTC-FCIQMC (all electrons and all virtual orbitals included) and xTC-CASCI$(10e, 8o)$ using the SA-CASSCF-Jastrow factor. As expected, the non-TC SA-CASSCF is not able to sufficiently capture all correlation effects of the excited states, resulting in highly overestimated excitation energies. The xTC CAS-only calculations improve on these, suggesting that the Jastrow factor is able to differentially capture some of the dynamical correlation missing in the SA-CASSCF calculation. However, since the resulting excitation energies are still largely outside chemical accuracy, some additional correlation via virtual orbitals needs to be included to achieve chemical accuracy.

Results are shown in figure 5.8. As shown, non-TC SA-CASSCF fails to capture all correlation effects of the excited states and tends to overestimate the excitation energies. The xTC CAS-only calculations are better, although still far from desired accuracy, and not nearly as accurate as the all-electron xTC calculations. However, since

the diagonalisation is performed only in a small active space, this is still a noteworthy improvement.

5.6. Conclusion and Outlook

We have presented a framework for using the transcorrelated method for solving problems of strongly multireference character. We find challenging problems like the nitrogen dissociation curve may be solved by modifying the ansatz for Φ with a multireference CI expansion, and appropriately changing the corresponding 1RDM for use in the xTC approximation. Particularly effective choices for Φ are FCIQMC and CASSCF. Using a small FCIQMC calculation has already been found to effectively solve the issues experienced with a HF ansatz, and its use with effective core potentials (ECPs) to efficienctly describe the core region with TC has already been explored.[321]

Furthermore, with this methodology we can tailor our TC Hamiltonian to target specific states of interest. Using state-averaged CASSCF, we find that this approach yields accurate results for excitation energies of the nitrogen dimer. Moreoever, we show that by combining the TC method to capture dynamical correlation with CASSCF to capture static correlation, we have improved accuracy. However, results suggest we may need larger active spaces and/or a transcorrelated perturbative treatment such as xTC-CASPT2.

While this approach has already proven remarkably effective for a few systems, one key outlook is to study a wider array of complex problems, including transition metals and periodic solids, in addition to a larger selection of smaller molecules beyond N_2. We may also wish to continue the procedure self-consistently. That is, we may use the resulting right-eigenvector of $\hat{H}_{\mathrm{TC}}$ as the ansatz for Φ in a subsequent calculation (as well as use the 1RDM, which may now be non-Hermitian), and continue until some convergence is achieved. Similarly, this study is a promising start for a TC-MCSCF method, wherein the orbitals, CI coefficients, and Jastrow factor parameters are all optimised simultaneously with a self-consistent algorithm, and no all-electron TC calculation is required or, as mentioned, we might add a transcorrelated perturbative correction as in CASPT2.

Minimal and Modular Jastrow Factors for the Transcorrelated Method

The contents of this chapter are planned to be expanded for a future publication. Some contents may be repeated there.

6.1. Introduction

So far, the TC methods we have discussed had one major bottleneck in common: the optimisation of the Jastrow factor. While VMC in itself does not scale unfavourably, the fact that we need so many VMC cycles in order to properly optimise for TC (see chapter 4) results in a significant computational cost. For large systems, this can become the prohibitive step in the workflow.

Here we explore some alternative avenues to construct Jastrow factors for use in the transcorrelated method. In particular, we consider constructing especially simple (minimal), or "universal" Jastrow factors that do not need any optimisation. These have already been introduced in the literature,[216,322–324] and are constructed to satisfy cusp conditions. We will also construct Jastrow factors from those optimised for atomic systems. That is, we optimise the Jastrow factor for the atom, and use the same Jastrow factor for molecules (so that the molecular Jastrow is the sum of atomic Jastrows). In effect, this would allow us to compile sophisticated atomic Jastrow factors that may be stored in a database for easy retrieval when considering larger systems, without any optimisation. We also consider keeping some components of these Jastrows fixed, while optimising other elements as a "molecular correction" to the atomic Jastrow factor.

6.2. Theory

In this chapter, we study various choices of Jastrow factors to avoid the need for lengthy optimisation. These may be divided into two broad categories: minimal Jastrow factors and atomic Jastrow factors.

6.2.1. Minimal Jastrow Factors

A minimal form for the Jastrow factor has already been presented by Fournais *et al* in 2005,[323] and has occasionally made further appearances in the literature.[216,324] The key advantage to using these is that we require no optimisation at all, while key disadvantages are that we must be careful about cut off functions, and we do not have as much flexibility so we cannot tailor the Jastrow factor to e.g. excited states.

One form of the Jastrow factor that might be called "universal" (that is, one that does not contain parameters that need optimisation per system) is given by[322,323]

$$J = -\sum_I \sum_i Z_I r_{iI} + \frac{1}{2} \sum_{i<j} r_{ij} + \frac{2-\pi}{6\pi} \sum_I \sum_{i<j} Z_I \boldsymbol{r}_{iI} \cdot \boldsymbol{r}_{jI} \ln\!\left(r_{iI}^2 + r_{jI}^2\right), \quad (6.1)$$

where, as usual, upper case indices denote the nucleus, and lower case indices represent electrons. The first term resolves the electron-nucleus cusps, the second term resolves the electron-electron cusps, and finally the last term resolves electron-electron-nucleus cusps. However, these terms are unbounded, and indeed are valid only close to coalescence points. For this reason, we introduce cutoff functions on each term.

Using the form of the cutoff functions used in previous chapters, that is $t(r, L) = (1 - r/L)^3 \Theta(r - L)$, we find unreasonable energies even for extremely small cutoffs. For example, with this form of cutoff with the electron-nucleus and electron-electron terms, and $L = 0.1$ Bohr, we get a reference energy of -246.604 for N_2 with the aug-cc-pVTZ basis set, whereas the HEAT result is known to be -109.5425.[285]

Thus, we instead use gaussian cutoffs, which are smooth but still decay rapidly for large distances. Our Jastrow factor therefore becomes

$$J = -\sum_I \sum_i Z_I r_{iI} e^{-r_{iI}^2/L_{en}^2} + \frac{1}{2} \sum_{i<j} r_{ij} e^{-r_{ij}^2/L_{ee}^2}$$
$$+ \frac{2-\pi}{6\pi} \sum_I \sum_{i<j} Z_I \boldsymbol{r}_{iI} \cdot \boldsymbol{r}_{jI} \ln\!\left(r_{iI}^2 + r_{jI}^2\right) e^{-r_{iI}^2/L_{een}^2} e^{-r_{jI}^2/L_{een}^2}, \quad (6.2)$$

where L_{ee}, L_{en}, L_{een} are cutoff parameters. For this study, we take $L := L_{ee} = L_{en} = L_{een}$, and consider three variants of equation 6.2:

– The full analytical Jastrow factor, as in equation 6.2. We shall dub this the "Fournais-Jastrow" factor.

– Neglecting the electron-electron-nucleus term and resolving the electron-nucleus cusp using the approach described in chapter 4 (instead of via the r_{iI} terms). We'll dub this the $ee + en$-Jastrow.

– Neglecting both the electron-electron-nucleus and electron-nucleus terms. This has already been studied in the context of TC-DMRG,[216] though the choice of cutoffs

may result in uncontrollably nonvariational energies. This is the simplest form, containing only r_{ij} terms, and we dub it the "*ee*-Jastrow".

Consider again the nitrogen molecule at equilibrium. We treat the result from HEAT[285] as the exact nonrelativistic ground state energy, and so this is as a lower bound for the lowest eigenvalue of $\hat{H}_{\text{TC}}$ for the various Jastrow factors. For each of the three choices above, there is only one parameter, L. For the Fournais-Jastrow factor, we find that the reference xTC-energy with a cutoff of $L = 0.1$ Bohr at aug-cc-pVTZ is -115.122 Hartree, which is below the HEAT result. This may be due to numerical issues from the complexity of this Jastrow factor form (none of the other ones we have explored included a dot product, for example). It may also indicate a more complex relationship between the TC energy and the cutoff L. As shown in chapter 4, the TC energy is sensitive to changes in the linear Jastrow factor parameters. As this term appears nonlinearly, determining L may be more challenging. However, for vanishingly small L, the TC and non-TC energies should be roughly equal. Even at $L = 0.1$ Bohr this is not the case, implying that for the Fournais-Jastrow factor, the TC energy is particularly sensitive to the choice of cutoff. Therefore, for simplicity, we exclude the Fournais-Jastrow factor from this study and consider only the *ee*- and *ee + en*-Jastrow factors.

The energy of N_2 and N with the aug-cc-pVTZ basis set for the *ee*- and *ee + en*-Jastrow factors are shown in figure 6.1. Shown are three non-TC energies: the canonical (RHF) reference energy, the canonical FCIQMC energy, and the HEAT (effectively CBS) energy. Plotted as a function of L are three TC energies, the xTC (RHF) reference energy, the xTC-MP2 energy, and the xTC-FCIQMC energy. For small L, we expect the non-TC- and xTC- reference and FCIQMC energies to coincide, which they approximately do. In contrast, for large L we expect the energies to become unstable, as we start to include spurious long-range correlation. However, it is worth noting that according to these plots, "long-range" is already at ≈ 0.4 Bohr, as the energies are all below the HEAT result. In every case, $L = 0.3$ Bohr is shown to result in energies above that of HEAT but below that of non-TC. We therefore use $L = 0.3$ Bohr as our "universal" cutoff for these Jastrow factors.

Since TC amounts to a similarity transformation, and any similarity transformation exactly preserves the eigenspectrum in the CBS limit, we know that this undesirable behaviour must be a basis set effect. Moreover, since the effects of these Jastrow factors are extremely localised near the nuclei, we might expect core-valence basis functions to aid in the TC energies. Figure 6.2 shows the effect of adding core functions to the basis set for the *ee + en*-Jastrow factor. We see that the core functions increase the (positive) correlation energies to bring it closer to the CBS limit, as does increasing the size (cardinal number) of the basis set. However, even cc-pCVQZ is not enough to bring the xTC-MP2 energy above the HEAT result for cutoff values above 0.4 Bohr. This suggests a very strong basis set effect, and merits further studies.
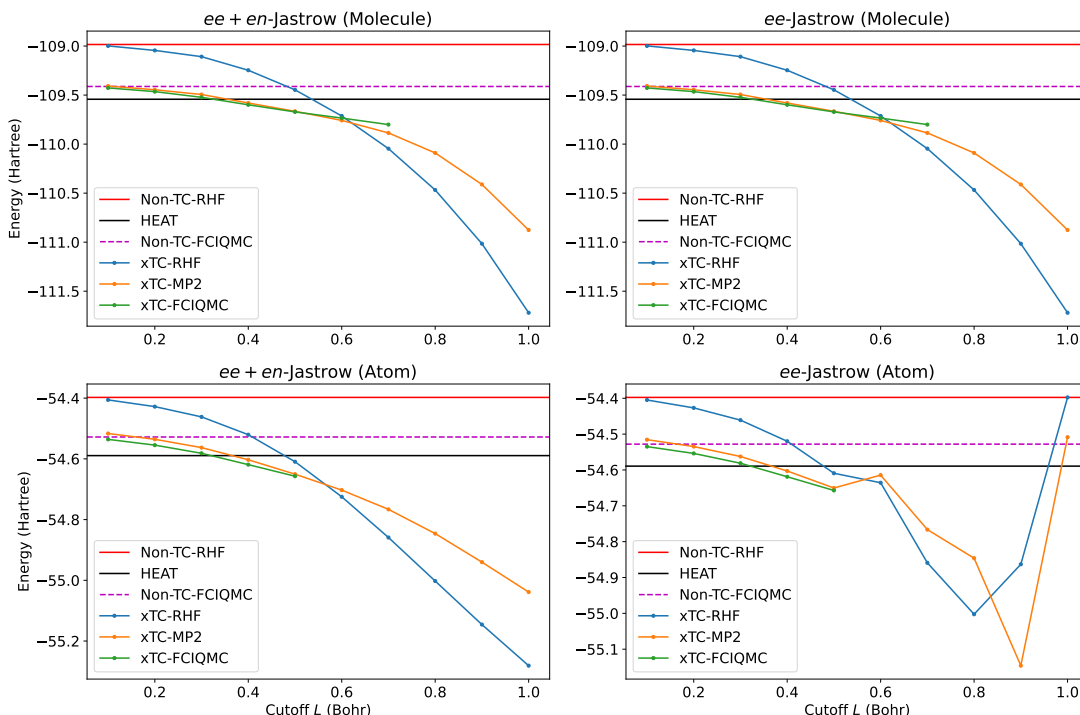
Figure 6.1. Energy estimates for the nitrogen molecule (top panel) and atom (bottom panel) as a function of a cutoff parameter L for the $ee + en$- (left panel) and ee-Jastrow (right panel) factors. The non-TC reference (RHF) energies are represented by a horizontal red line, the HEAT result by a horizontal black line, and the non-TC-FCIQMC by a dashed line. Since HEAT is a CBS-extrapolated method, and variational theories are typically preferred, we aim to keep our TC energies above this value. However, we also expect TC to improve upon the canonical FCIQMC energy. We therefore take the value of L that leads to energies above the HEAT result, which is $L = 0.3$ for all four plots. Large- and small-L limits behave as expected, with the former giving undesirable results and the latter approximating the non-TC results in that basis set. All calculations were performed with the aug-cc-pVTZ basis set. Missing xTC-FCIQMC points are due to numerical instability caused by a positive correlation energy. In practice, these may be resolved by setting a positive shift, but these results are anyway undesirable. Noise caused by the large cutoff in the ee-Jastrow factor for the atom likely indicates unfavourable changes to the one-electron density encoded in the HF orbitals. Adding an electron-nucleus term has been known to give a stabilising effect counteracting this,[170,253], which is consistent with the fact that the $ee + en$-Jastrow factor does not exhibit this noise.

6.2.2. Atomic Jastrow Factors

Another approach we consider in reducing the need for optimisation is to optimise Jastrow factors for atomic systems and then reuse them in the molecular context. This can be considered analogous to mixing atomic orbitals orbitals in order to calculate molecular orbitals. The key advantage here is flexibility while the key disadvantages include the need to optimise for the atoms, and the need to make a choice for the form of the Jastrow factor. In principle, we could produce a database of sophisticated atomic Jastrow factors
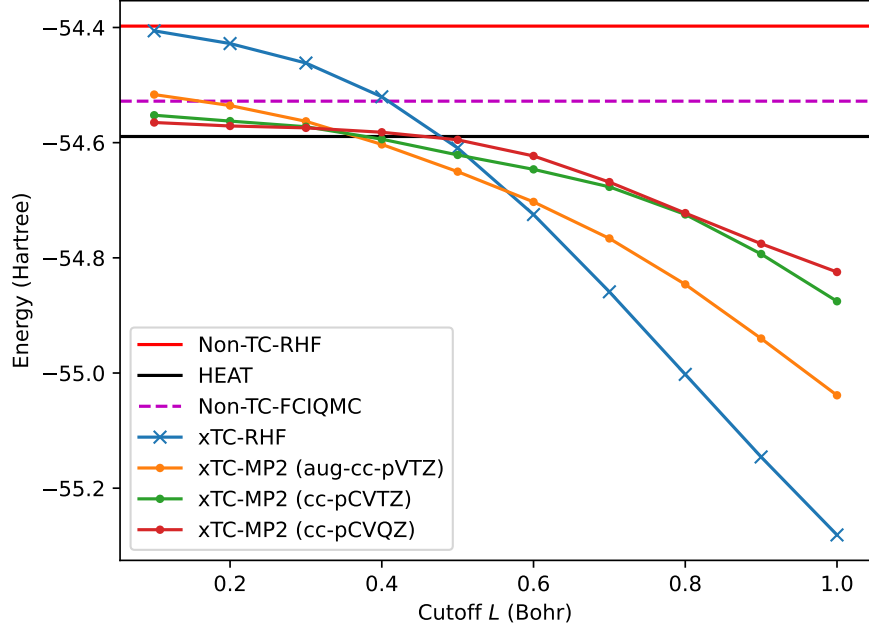
Figure 6.2. MP2 energies as a function of the cutoff L for the $ee + en$-Jastrow factor. The xTC-RHF values are roughly the same for each basis set shown. Adding core functions significantly improves the curve to be closer to the HEAT result, resulting in larger (positive) correlation energies. Increasing the cardinal number of the basis set also improves the curve, but the effect is less pronounced. Nevertheless, these results indicate a strong basis set effect in the cutoff, particularly for core functions. Adding diffuse functions have negligible effects on the curve (e.g. xTC-MP2 curves for aug-cc-pCVTZ, not shown here, and cc-pCVTZ roughly coincide).

that can then be queried to construct Jastrow factors for molecules or periodic systems. We use the Jastrow forms considered earlier in this dissertation,

$$J = \sum_{i<j}^{N} v(r_{ij}) + \sum_{i}^{N} \sum_{I}^{N_A} \chi(r_{iI}) + \sum_{i<j}^{N} \sum_{I}^{N_A} f(r_{ij}, r_{iI}, r_{jI}), \tag{6.3}$$

with

$$v(r_{ij}) = t(r_{ij}, L_v) \sum_{k} a_k r_{ij}^k, \tag{6.4}$$

$$\chi(r_{iI}) = t(r_{iI}, L_\chi) \sum_{k} b_k r_{iI}^k, \tag{6.5}$$

$$f(r_{ij}, r_i, r_j) = t(r_{iI}, L_f) t(r_{jI}, L_f) \sum_{k,l,m} c_{klm} r_{ij}^k r_{iI}^l r_{jI}^m, \tag{6.6}$$

and the same cutoff functions $t(r, L) = (1 - r/L)^3 \Theta(r - L)$. However, we do not want to include long-range (with respect to the nucleus) correlation in the atomic Jastrow factors, since this may bias the molecular calculations. We therefore use $L_f = L_\chi = 1.0$ Bohr and $L_v = 4.5$ Bohr. We consider the following variants:

– The Jastrow factor is kept constant for the molecule, i.e. we simply use the atomic Jastrow factors as they are. We refer to these as simply "atomic" Jastrow factors.

– We optimise the atomic Jastrow factor using all terms. Then, when calculating molecules, we fix all terms involving nuclei and re-optimise the electron-electron terms. We refer to these as "atomic+ee" Jastrow factors.

We find these to yield results above HEAT and therefore do not concern ourselves with a cutoff analysis as in the case of the minimal Jastrow factors.

6.3. Results

6.3.1. Computational Details

We revisit the nitrogen binding curve from chapter 5 as a stress test for our Jastrow factors, as well as the atomisation of the molecules N_2, C_2, O_2, and CN from chapter 4. As before, we compare the results of the different choices of Jastrow factors against experiment[307] and the atomisation energies to HEAT.[285] The binding curve is calculated with the aug-cc-pVTZ basis set, while the atomisation energies are calculated with the aug-cc-pVXZ basis sets for $X = D, T, Q$.

For the multideterminantal optimisation, we use a small FCIQMC calculation as it was found to perform particularly well in chapter 5. Non-TC HF and CASCI calculations were performed using `pyscf`,[289] VMC optimisation was performed using `CASINO`,[253] FCIQMC calculations using `NECI`,[267] and MRCI-F12 calculations for comparison using `Molpro`.[311–313]

6.3.2. Atomisation Energies

For the atomisation energies, we focus on the $ee + en$ and atomic+ee Jastrow factors, as these are the most "complete" of each category of Jastrow factors presented here. The absolute xTC-FCIQMC energy estimates using the various methods for the systems considered in this chapter are presented in table 6.1. Also presented there are the non-TC FCIQMC and TC-FCIQMC (that is, with no additional approximations such as xTC), as presented in chapter 4.

From this table, it is clear that all the TC methods have much more rapid convergence toward the CBS limit. The "full TC" method, as presented in chapter 4 which uses the form of equation 6.3 with larger cutoff values (4.5, 4 and 4 Bohr for the electron-electron, electron-nucleus and electron-electron-nucleus terms, respectively) converges to the CBS limit most rapidly, as expected. The atomic+ee total energies also converge to the CBS limit much more rapidly than conventional FCIQMC, and is remarkably close at cc-pVQZ. While the absolute energies for the molecules still differ from the CBS limit by up to

System	Method	cc-pVDZ	cc-pVTZ	cc-pVQZ	CBS[285,286]
N$_2$	Non-TC	-109.2809	-109.4014	-109.4653	
	Full TC	-109.4727	-109.5312	-109.5428	-109.5425
	$ee + en$	-109.4125	-109.5141	-109.5580	
	Atomic+ee	-109.4364	-109.5168	-109.5410	
C$_2$	Non-TC	-75.7320	-75.8094	-75.8578	-75.9240
	Full TC	-75.8844	-75.9197	-75.9272	
	$ee + en$	-75.8153	-75.8823	-75.9248	-75.9265
	Atomic+ee	-75.8567	-75.9082	-75.9248	
O$_2$	Non-TC	-149.9915	-150.1554	-150.2362	
	Full TC	-150.2216	-150.3078	-150.3244	-150.3273
	$ee + en$	-150.1972	-150.3243	-150.3662	
	Atomic+ee	-150.1763	-150.2908	-150.3221	
CN	Non-TC	-92.4970	-92.5954	-92.6517	
	Full TC	-92.6671	-92.7152	-92.7247	-92.7232
	$ee + en$	-92.6039	-92.6872	-92.7281	
	Atomic+ee	-92.6331	-92.6992	-92.7205	
N	Non-TC	-54.4801	-54.5252	-54.5535	
	Full TC	-54.5622	-54.5842	-54.5896	-54.5893
	$ee + en$	-54.5425	-54.5792	-54.5986	
	Atomic+ee	-54.5477	-54.5786	-54.5890	
C	Non-TC	-37.7619	-37.7900	-37.8126	
	Full TC	-37.8293	-37.8427	-37.8459	-37.8450
	$ee + en$	-37.8020	-37.8256	-37.8437	
	Atomic+ee	-37.8193	-37.8374	-37.8451	
O	Non-TC	-74.9117	-74.9853	-75.0236	
	Full TC	-75.0226	-75.0572	-75.0665	-75.0674
	$ee + en$	-75.0123	-75.0686	-75.0882	
	Atomic+ee	-74.9986	-75.0514	-75.0657	

Table 6.1. Absolute energies for N$_2$, C$_2$, O$_2$, CN, N, O, and C using conventional non-TC-FCIQMC, the full TC workflow with larger cutoffs (see chapter 4), the minimal $ee+en$ Jastrow factor with $L = 0.3$ Bohr, and the atomic+ee Jastrow factor, where the Jastrow factors for the atoms are reused for the molecule, and the ee term is reoptimised with the RHF molecular wave function ansatz. Values are reported in Hartree for the cc-pVDZ, cc-pVTZ and cc-pVQZ basis sets and compared against high-accuracy benchmarks.[285,286] We find that all TC methods converge rapidly to the CBS limit, but the $ee+en$ Jastrow factor has a tendency to "overshoot" the limit at cc-pVQZ, suggesting a nonlinear convergence. Values in red are below the CBS limit. This only occurs at cc-pVQZ with the minimal Jastrow factor; the atomic Jastrow factor is more stable, and is the only TC method presented to consistently be above the energy at the CBS limit.

about 5 mHa in the cc-pVQZ basis, all atom calculations are within 1 mHa of the CBS limit.

The (minimal) $ee + en$-Jastrow factor also exhibits rapid basis set convergence, but unlike the other methods also results in energies below the CBS limit with the cc-pVQZ basis set. These discrepencies can be quite large, almost 40 mHa in the case of CN. This might suggest that our "universal" cutoff proposed in figure 6.1 really is not so "universal" after all. The L chosen was not properly optimised, and was determined using only one system, so it is not wholly surprising that we might get questionable TC final energies using it.

However, in practice we are typically much more interested in relative energies. The atomisation energies for these molecules, obtained from the values in table 6.1 are shown in table 6.2 and plotted in figure 6.3. We find that despite the nonvariational energies found in table 6.1, error cancellation when using the $ee + en$ terms are enough to give sensible atomisation energy estimates. However, they are consistently outperformed by the atomic+ee Jastrow factor, which is also a relatively inexpensive calculation due to the lack of additional terms needed to be optimised. This data suggests to use Jastrow factors based on atomic systems for systems that are not amenable to VMC calculations for all parameters.

As with the total energies, the TC atomisation energies converge much faster with respect to basis-set size in comparison to the non-TC calculations. While not quite on par with the other TC methods, it appears to converge faster with respect to basis set size than the non-TC energies. Nevertheless, the uncontrolled nonvariational error arising from a fixed L value suggests caution when considering the use of such minimal Jastrow factors. In contrast, the atomic+ee is much more robust.

Unlike with full TC, the atomic and minimal Jastrow factors presented do not consistently reach chemical accuracy (defined as within 1.6 mHa of the value at the CBS limit) for the atomisation of any molecules. However, they are still considerably closer than non-TC. Moreover, the atomic Jastrow factor consistently outperforms the minimal form. Since, as discussed, this form also gives absolute energies above the CBS limit, converging from above, this is strong evidence that the atomic+ee Jastrow factor form is the preferred method for systems that are too large for optimising all parameters simultaneously as done in previous chapters. While we have reduced accuracy, the trade off of not needing to calculate all parameters may be worthwhile for future applications of the TC method. In the full TC calculations, the Jastrow factors had 44 optimisable parameters; in the atomic+ee case, there are only 8 optimisable parameters.

6.3.3. Binding Curves

Here we report the N_2 binding curve with the aug-cc-pVTZ basis set for the various choices of Jastrow factors. Values for the interatomic distance range between 1.92 and 10 Bohr.
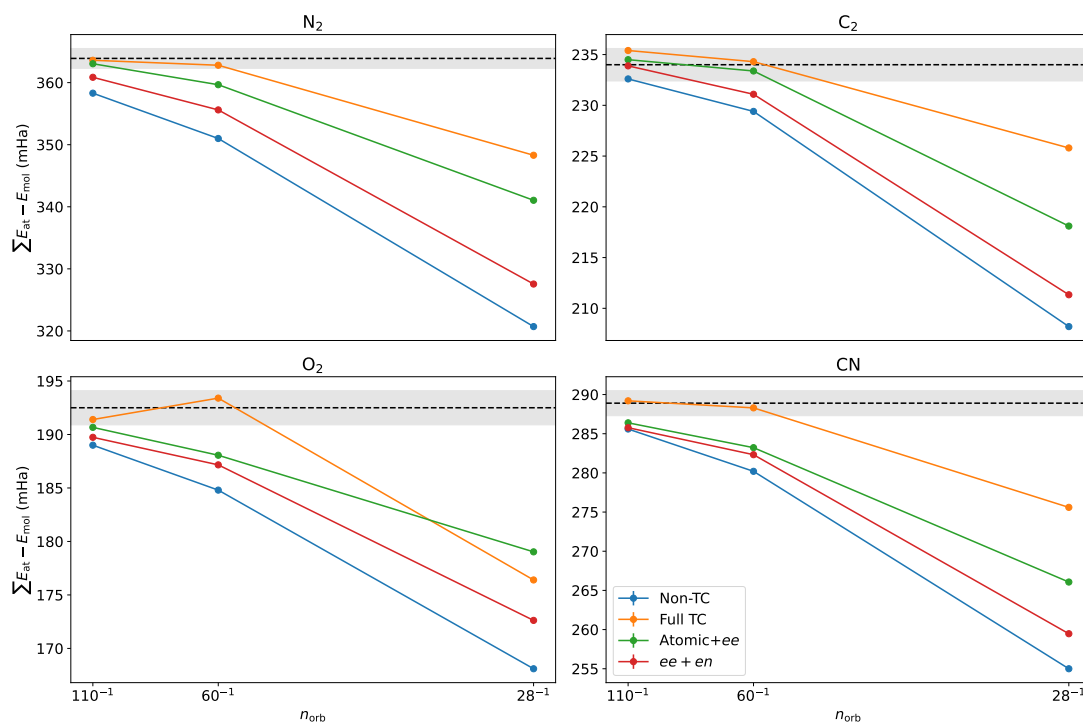
Figure 6.3. Atomisation energies for the molecules N_2, O_2, C_2 and CN as a function of the reciprocal of the number of orbitals. The CBS limit is represented by a dashed line and a grey region denotes the area within chemical accuracy (± 1.6 mHa) of this value. From these plots, we can see a consistent trend where the full TC treatment from chapter 4 performs best in that it converges most rapidly to the CBS, with the atomic+ee Jastrow factor approach also performing well. Interestingly, the $ee + en$ Jastrow factor approach also performs well, despite the undesirable absolute energies discussed in table 6.1, suggesting considerable error cancellation. In terms of atomisation energies, this approach also performs reasonably well, though the only data point within chemical accuracy is C_2 in the cc-pVQZ basis, but even non-TC is chemically accurate in this case.

Molecule	Method	cc-pVDZ	cc-pVTZ	cc-pVQZ	CBS[285]
N_2	Non-TC	320.7	351.0	358.3	
	Full TC	348.3	**362.8**	**363.6**	363.9
	ee + en	327.6	355.6	360.9	
	Atomic+*ee*	341.0	359.7	**363.0**	
C_2	Non-TC	208.2	229.4	**232.6**	234.0
	Full TC	225.8	**234.3**	235.4	
	ee + en	211.3	231.1	**233.9**	236.5
	Atomic+*ee*	218.1	**233.4**	**234.5**	
O_2	Non-TC	168.1	184.8	189.0	
	Full TC	176.4	**193.4**	**191.4**	192.5
	ee + en	172.6	187.2	189.7	
	Atomic+*ee*	179.0	188.1	190.7	
CN	Non-TC	255.0	280.2	285.6	
	Full TC	275.6	**288.3**	**289.2**	288.9
	ee + en	259.5	282.3	285.8	
	Atomic+*ee*	266.1	283.2	286.4	

Table 6.2. Atomisation energies for N_2, C_2, O_2 and CN. Results are shown for the *ee+en* and atomic+*ee* Jastrow factors, and compared to non-TC and full TC for the cc-pVDZ, cc-pVTZ and cc-pVQZ basis sets, as well as to the CBS limit. Bold face indicates being within chemical accuracy of one of the benchmark CBS values.

The Jastrow factors used were the atomic Jastrow factor, the atomic Jastrow factor with electron-electron terms optimised for the molecule (using a RHF or FCIQMC ansatz), and the minimal *ee*- and *ee + en*-Jastrow factors. The binding curve for the FCIQMC-Jastrow factor (i.e. multideterminant ansatz with all terms optimised) from chapter 5 is also shown for comparison. All choices of Jastrow factor give qualitatively correct binding curves. Plots for the binding curves are shown in figure 6.4, focusing on the minimum and dissociation regions. Compared to the FCIQMC-Jastrow factor of chapter 5, the binding curves using the Jastrow factors presented in this chapter shift the minimum upwards. This is to be expected, since we have fewer (or even no) parameters being optimised, and thus less flexibility in the VMC optimisation. Note that, while the cutoff L for the minimal Jastrow factors is not optimised, its value was determined in section 6.2.1 based on this same system at equilibrium. Based on the results in section 6.3.2, it is likely that this cutoff analysis effectively behaved as a surrogate for optimisation, and that $L = 0.3$ Bohr would not be optimal for other systems, and may even result in values below the CBS limit. This suggests that simple Jastrow factors such as the *ee*-Jastrow factor are viable, but that they need to be used carefully, and may need to be optimised for each system.

It is also worth noting that unlike the RHF-Jastrow factor in chapter 5, the single-reference electron-electron optimisation with the atomic Jastrow factor performs well, and even gives the correct behaviour in the dissociated limit. This may be because the atomic Jastrow factor is already a good approximation for the dissociated limit,
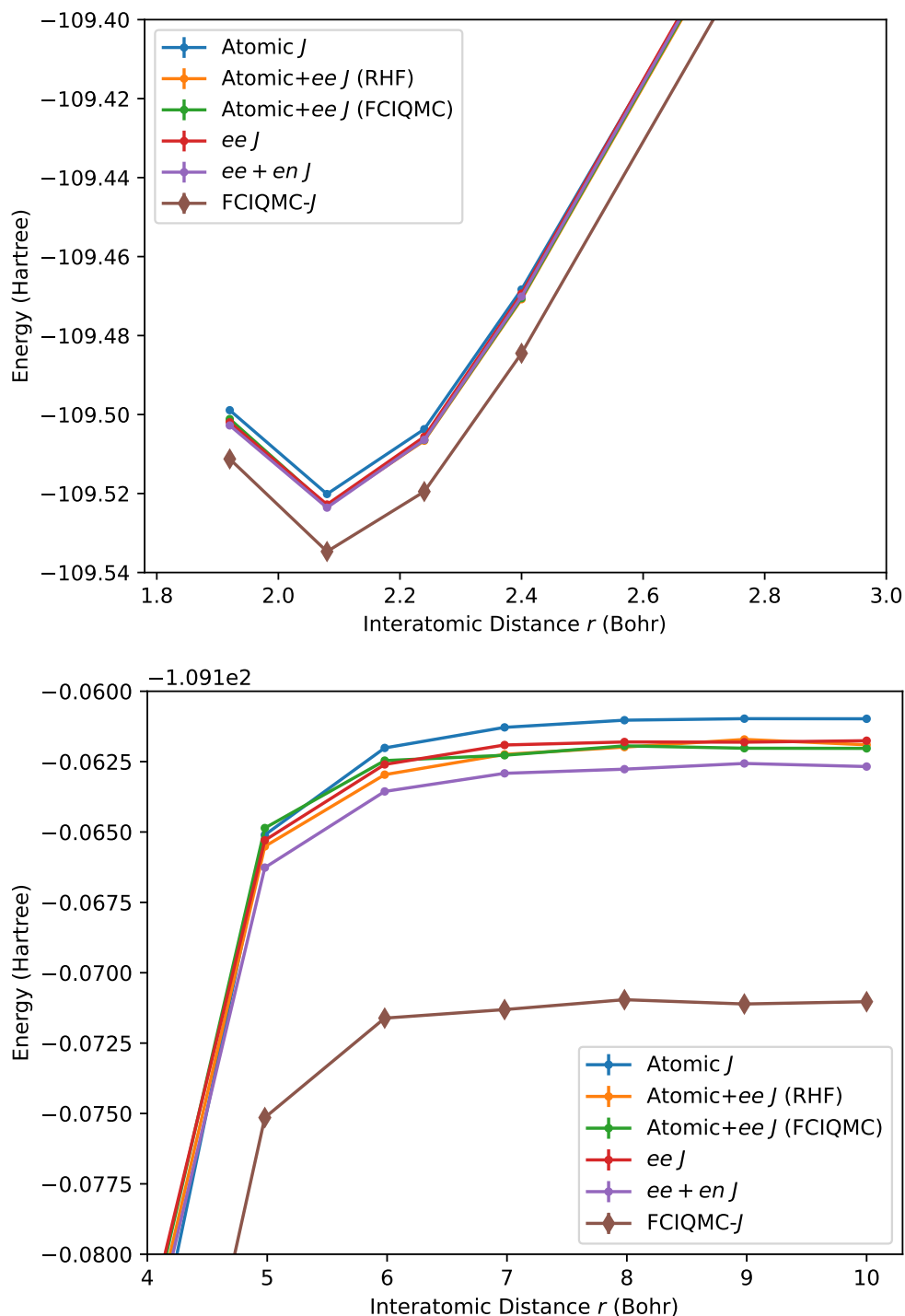
Figure 6.4. Zoomed in plots of the N_2 binding curves for the various Jastrow factor choices. All forms give qualitatively correct binding curves. Included for comparison is the FCIQMC-Jastrow from chapter 5, which was shown to be good both at equilibrium and dissociation. Compared to this, all Jastrow factors newly presented in this chapter shift the minimum upwards, with all of them having similar values in this region except for the pure atomic Jastrow factor. This is likely because the eleectron-electron terms are suboptimal. Slight noise in the dissociated limit are likely connected to the error in the VMC optimisation, which was shown in chapter 4 to be on the order of 0.1 mHa.

and so the electron-electron optimisation in the molecule does little extra work. It also does not have as much flexibility as the RHF-Jastrow factor from chapter 5, so the poor TC ansatz cannot do as much harm, though it is responsible for a higher energy estimate in the dissociated limit, as can be seen by comparing against the multireference (FCIQMC)-optimised electron-electron Jastrow factor.
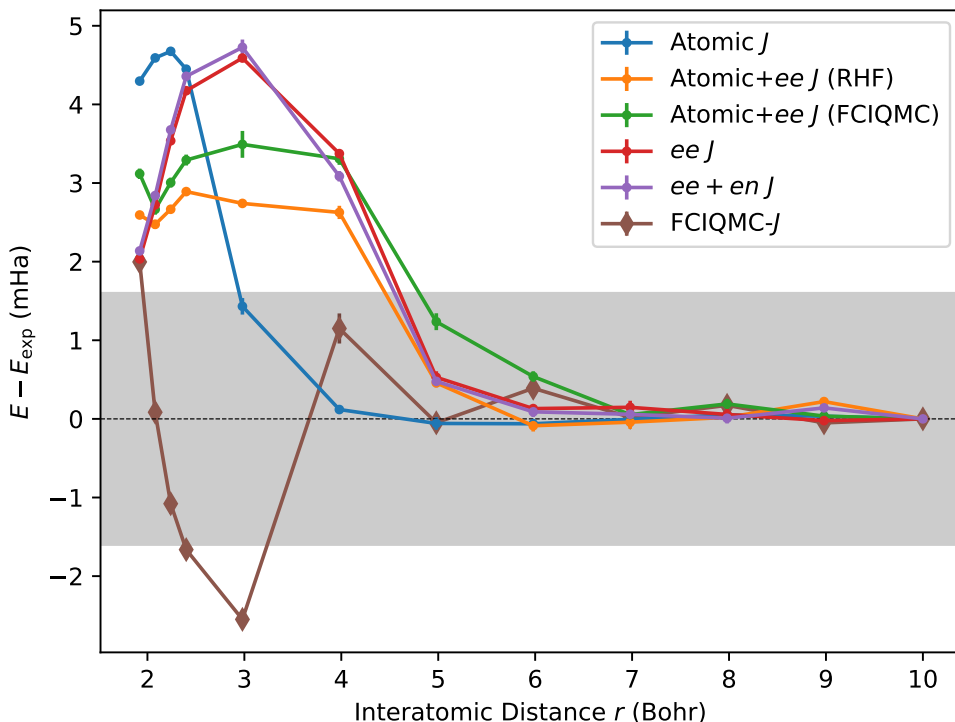
Figure 6.5. The binding curves relative to a fit to experimental data[307] for the nitrogen dimer, calculated in the aug-cc-pVTZ basis for various Jastrow factor forms. All curves are normalised such that the value at 10 Bohr is 0 Hartree. The grey shaded area denotes a region of ± 1.6 mHa (i.e. chemical accuracy). For reference, the multideterminantal FCIQMC-Jastrow from chapter 5 is shown in brown (with diamonds). Both minimal Jastrow factor forms (ee and $ee + en$, denoting with only electron-electron or with both electron-electron and electron-nucleus terms, respectively) stay above the experimental result, and are indeed very close to each other. The atomic Jastrow factors, whether containing ee terms optimised for the molecule (RHF or FCIQMC as the Φ ansatz) or not optimised for the molecule (atomic J), have a less smooth behaviour, but are generally closer to the experimental result.

The energy estimates relative to experiment (normalised such that the energy at 10 Bohr is 0 Hartree) is shown in figure 6.5. In contrast to the multireference Jastrow factors presented in chapter 5, such as the FCIQMC-Jastrow factor reproduced here, the atomic and minimal Jastrow factors stay above the experimental result and converge relatively monotonically. That said, they are overall not as close to the experimental result as the

FCIQMC-Jastrow factor across the binding curve. It is also worth noting that the *ee*- and *ee + en*-Jastrow factors result in very similar curves, suggesting only a minimal effect from the nuclear term.

Since all of these Jastrow factors are strongly localised, we expect them all to be well-behaved at long interatomic separation, i.e. they should be size consistent, and that is consistent with figure 5.6. The possible exception is the single-determinant-optimised atomic+*ee* Jastrow factor, as the electron-electron term is optimised with a non-size-consistent wave function ansatz. Howeer, as observed even this curve appears well-behaved. We approximate the size consistency error by the difference in energies between the molecule and the sum of its parts and show the results in table 6.3. Indeed, we find all Jastrow factors to be size consistent, comparable to MRCI-D-F12.

Jastrow Factor	$\sum E_{\text{atom}} - E_{\text{molecule}}(r = 10)$ (mHa)
Atomic	-0.2(0)
Atomic+*ee* (RHF)	0.5(0)
Atomic+*ee* (FCIQMC)	0.9(0)
ee	-0.6(0)
ee + en	-0.6(1)
MRCI-D-F12	-0.9

Table 6.3. Size consistency error, measured as the difference between the sum of energies for the constituent atoms in isolation and the energy of the molecule at dissocation (taken to be 10 Bohr). We see that all methods presented here have size consistency errors that are comparable to MRCI-D-F12.

Finally, the atomisation energy estimates using the binding curves, i.e. the difference between the energy in the dissociated limit and that at equilibrium, are shown in table 6.4. All Jastrow factors perform similarly well, comparable to MRCI-D-F12. While not as accurate as previous methods presented in this dissertation, these approaches have fewer parameters to optimise.

6.4. Conclusion and Outlook

We have presented an alternative approach to constructing Jastrow factors, either using a minimal approach focusing on describing analytical behaviour near coalescence points, or reusing Jastrow factors optimised for small systems (such as an atom) for larger systems (such as a molecule). We have shown that the resulting Jastrow factors are size consistent and provide rapid basis set convergence when compared against conventional electronic structure methods. These updated workflows represent significant improvements in the scalability and ease of use for TC methods, while arguably being more conceptually satisfying by making the Jastrow factors more general.

It is, however, worth pointing out some shortcomings of these approaches. Perhaps the most obvious shortcoming is the relative lack of flexibility in the Jastrow factor, as

Jastrow Factor	Atomisation Energy (mHa)
Atomic	358.9
Atomic+ee (RHF)	361.8
Atomic+ee (FCIQMC)	361.7
ee	361.0
$ee + en$	360.3
MRCI-D-F12	359.5
HEAT[285]	363.9
Experiment[307]	363.7

Table 6.4. Atomisation energies for the nitrogen dimer using only the binding curve. We see that all methods presented here have atomisation energies that are comparable to MRCI-D-F12, with most slightly outperforming it. While not within chemical accuracy like full TC or the multireference Jastrows presented in chapter 5, these results are still promising on account of their relative simplicity.

either there are no terms to optimise or relatively few compared to previous approaches presented in this dissertation. Moreoever, with the exception of the atomic+ee Jastrow factor with the electron-electron term optimised using an FCIQMC ansatz, there is no clear way for these methods to be tailored for specific states. Additionally, the use of minimal Jastrow factors may lead to TC energies well below the CBS limit if not careful in the choice of the cutoff function. While these appear to be mostly compensated for by error cancellation when computing atomisation energies, the nonmonotonic convergence to the CBS limit is nevertheless undesirable. This suggests future studies on the effects of the cutoffs and the potential need for optimising its parameters for each state. In this sense, the form of the minimal Jastrow factor may be "universal", but its cutoffs may not be. Since this convergence appears strongly linked to the core region, it may also be worthwhile studying the effects of using futher core functions, or ECPs, as well as a more careful introduction of the electron-electron-nucleus term in the Fournais-Jastrow factor.

These results suggest that a realistic approach to studying larger systems and/or performing optimisation deterministically, with the transcorrelated method would involve combining Jastrow factors optimised for small systems and reoptimising some terms (such as the electron-electron term) for the larger system, as a form of correction. This would allow for a more computationally tractable approach with fewer optimisable parameters, while still benefiting from rapid convergence to the CBS limit.

Summary and Outlook

Transcorrelation has recently enjoyed somewhat of a revival since its early development in the 1960s, particularly when combined with other methods such as CC, DMRG, or FCIQMC. The topic of this dissertation has been to further explore modern developments of the TC method, especially applied to FCIQMC. Both of these methods combine ideas from QMC and apply them to wave function methods. FCIQMC stochastically samples the (ground state) wave function of a system in analogy to DMC to give FCI-accuracy. The TC method adopts the Jastrow factor typically encountered in VMC to correlate the electrons and explicitly take into account analytically-known behaviour at coalescence, a task difficult for conventional methods. Combining the two results in an essentially-exact stochastic method with rapid convergence towards the CBS limit.

The primarily contributions to this field represented in this dissertation are threefold. First, in chapter 4, using flexible Jastrow factors with VMC to minimise the variance of the reference energy is proposed and shown to result in not only highly-accurate energies with chemically-accurate atomisation energies with only the cc-pVTZ basis set (compared to cc-pV5Z for non-TC-FCI), but it also compactifies the wave function, making it better suited to sampling methods such as FCIQMC.

Second, in chapter 5 we have shown that the TC method was not able to accurately describe problems of strongly multireference character, such as the dissociation of N_2, despite using FCIQMC after Jastrow-factor optimisation. Modifying the TC ansatz to accommodate multireference wave functions, we show that we can still benefit from using the TC method even in such challenging problems. Moreover, this introduces additional flexibility in the optimisation procedure, allowing us to tailor our Jastrow factor for specific states. We showed that this can be used to determine excited state energies with high accuracy.

Finally, in chapter 6 we consider alternative forms for the Jastrow factor, either allowing much faster optimisation or bypassing it altogether, implying better suitability for larger systems. We considered two different categories of such Jastrow factors. In the first, the Jastrow factors were in a sense minimal, insofar as primarily handling known analytical behaviour of the wave function near coalescence points. These Jastrow factors were shown to sometimes result in nonvariational energies, but they did nevertheless exhibit favourable error cancellation. This is suggestive for future studies to prevent the nonvariational energies from occurring, possibly by optimising the cutoff range of the Jastrow factor

terms. The second category of Jastrow factors were the atomic Jastrow factors, wherein the Jastrow factors are optimised for each atom in the system, and then these Jastrow factors are used in the molecule, only reoptimising a small portion of the parameters. This still resulted in clearly-superior basis-set convergence compared to conventional methods, while not requiring as much computational effort as the full optimisation. Moreover, the energies we found were all above the CBS limit. In principle, this opens the possibility of optimising sophisticated Jastrow factors for atoms across the periodic table, storing them in a database, and querying it when needed in a calculation, in analogy to what is already done with basis sets.

The methods proposed in this thesis have already been applied in other studies, such as applying TC to larger atoms[325] or combining TC with ECPs (where the FCIQMC-Jastrow approach of chapter 5 can also be used).[321] Other potential avenues of research include applying the methodologies to larger systems, such as solids, or embedding with the TC method. Combining the optimisation procedure outlined in chapter 4 with the active space multireference approach in chapter 5, we could conceivably optimise biorthogonal orbitals in a transcorrelated CASSCF-like procedure for high accuracy within an active space. This might also be combined with simpified Jastrow factors such as those presented in chapter 6 and deterministic optimisation as in Appendix A to allow for more robust but scalable calculations.

Despite having its roots in the 1960s, the methodology is still in its infancy, and as a result the most productive step forward would likely be to simply apply the methodology to new problems of chemical interest. This would necessarily lead to new methods, computational techniques, and code optimisations.

List of Figures

List of Tables

Bibliography

(1)   Laughlin, R. B.; Pines, D. "The Theory of Everything". Proceedings of the National Academy of Sciences **2000**, *97*, 28–31.

(2)   Ryden, B., *Introduction to Cosmology*; Cambridge University Press: 2017.

(3)   Jones, L. O.; Mosquera, M. A.; Schatz, G. C.; Ratner, M. A. "Embedding Methods for Quantum Chemistry: Applications from Materials to Life Sciences". Journal of the American Chemical Society **2020**, *142*, 3281–3295.

(4)   Christlmaier, E. M. C.; Kats, D.; Alavi, A.; Usvyat, D. "Full Configuration Interaction Quantum Monte Carlo Treatment of Fragments Embedded in a Periodic Mean Field". The Journal of Chemical Physics **2022**, *156*, 154107.

(5)   Helgaker, T.; Jorgensen, P.; Olsen, J., *Molecular Electronic-Structure Theory*; John Wiley & Sons: 2014.

(6)   Born, M.; Oppenheimer, R. "Zur Quantentheorie Der Molekeln". Annalen der Physik **1927**, *389*, 457–484.

(7)   Hubbard, J. "Electron Correlations in Narrow Energy Bands". Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **1963**, *276*, 238–257.

(8)   Lieb, E. H.; Wu, F. Y. "Absence of Mott Transition in an Exact Solution of the Short-Range, One-Band Model in One Dimension". Physical Review Letters **1968**, *20*, 1445–1448.

(9)   Liebermann, N. *The FCIQMC Sign Problem in the Real-Space Hubbard Model*, Doctoral Thesis, 2023.

(10) Motta, M.; Ceperley, D. M.; Chan, G. K. L.; Gomez, J. A.; Gull, E.; Guo, S.; Jiménez-Hoyos, C. A.; Lan, T. N.; Li, J.; Ma, F.; Millis, A. J.; Prokof'ev, N. V.; Ray, U.; Scuseria, G. E.; Sorella, S.; Stoudenmire, E. M.; Sun, Q.; Tupitsyn, I. S.; White, S. R.; Zgid, D.; Zhang, S. "Towards the Solution of the Many-Electron Problem in Real Materials: Equation of State of the Hydrogen Chain with State-of-the-Art Many-Body Methods". Physical Review X **2017**, *7*, 031059.

(11) Motta, M.; Genovese, C.; Ma, F.; Cui, Z. H.; Sawaya, R.; Chan, G. K. L.; Chepiga, N.; Helms, P.; Jiménez-Hoyos, C.; Millis, A. J.; Ray, U.; Ronca, E.; Shi, H.; Sorella, S.; Stoudenmire, E. M.; White, S. R.; Zhang, S. "Ground-State Properties of the Hydrogen Chain: Dimerization, Insulator-to-Metal Transition, and Magnetic Phases". Physical Review X **2020**, *10*, 031058.

(12)   Sakurai, J. J.; Napolitano, J., *Modern Quantum Mechanics*; Cambridge University Press: 2017.

(13)   Ascher, U. M.; Greif, C., *A First Course in Numerical Methods*; SIAM: 2011.

(14)   Dunning Jr., T. H. "Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen". The Journal of Chemical Physics **1989**, *90*, 1007–1023.

(15)   Feller, D.; Peterson, K. A.; Grant Hill, J. "On the Effectiveness of CCSD(T) Complete Basis Set Extrapolations for Atomization Energies". The Journal of Chemical Physics **2011**, *135*, 044102.

(16)   Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Koch, H.; Olsen, J.; Wilson, A. K. "Basis-Set Convergence in Correlated Calculations on Ne, N2, and H2O". Chemical Physics Letters **1998**, *286*, 243–252.

(17)   Halkier, A.; Helgaker, T.; Jørgensen, P.; Klopper, W.; Olsen, J. "Basis-Set Convergence of the Energy in Molecular Hartree–Fock Calculations". Chemical Physics Letters **1999**, *302*, 437–446.

(18)   Helgaker, T.; Klopper, W.; Koch, H.; Noga, J. "Basis-Set Convergence of Correlated Calculations on Water". The Journal of Chemical Physics **1997**, *106*, 9639–9646.

(19)   Jensen, F. "The Basis Set Convergence of the Hartree–Fock Energy for H2". The Journal of Chemical Physics **1999**, *110*, 6601–6605.

(20)   Pansini, F. N. N.; Neto, A. C.; Varandas, A. J. C. "Extrapolation of Hartree–Fock and Multiconfiguration Self-Consistent-Field Energies to the Complete Basis Set Limit". Theoretical Chemistry Accounts **2016**, *135*, 261.

(21)   Peterson, K. A.; Woon, D. E.; Dunning Jr., T. H. "Benchmark Calculations with Correlated Molecular Wave Functions. IV. The Classical Barrier Height of the H+H2→H2+H Reaction". The Journal of Chemical Physics **1994**, *100*, 7410–7415.

(22)   Woon, D. E.; Dunning Jr., T. H. "Benchmark Calculations with Correlated Molecular Wave Functions. VI. Second Row A2 and First Row/Second Row AB Diatomic Molecules". The Journal of Chemical Physics **1994**, *101*, 8877–8893.

(23)   Boys, S. F. "Electronic Wave Functions - I. A General Method of Calculation for the Stationary States of Any Molecular System". Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **1950**, *200*, 542–554.

(24)   Szabo, A.; Ostlund, N. S., *Modern Quantum Chemistry: Introduction to Advanced Electronic Structure Theory*; Courier Corporation: 2012.

(25)   Hastie, T.; Tibshirani, R.; Friedman, J., *The Elements of Statistical Learning: Data Mining, Inference, and Prediction, Second Edition*; Springer Science & Business Media: 2009.

(26) Bayes, T.; Price, n. "LII. An Essay towards Solving a Problem in the Doctrine of Chances. By the Late Rev. Mr. Bayes, F. R. S. Communicated by Mr. Price, in a Letter to John Canton, A. M. F. R. S". Philosophical Transactions of the Royal Society of London **1997**, *53*, 370–418.

(27) Hartree, D. R. "The Wave Mechanics of an Atom with a Non-Coulomb Central Field. Part II. Some Results and Discussion". Mathematical Proceedings of the Cambridge Philosophical Society **1928**, *24*, 111–132.

(28) Fock, V. "Näherungsmethode Zur Lösung Des Quantenmechanischen Mehrkörperproblems". Zeitschrift fur Physik **1930**, *61*, 126–148.

(29) Slater, J. C. "Note on Hartree's Method". Physical Review **1930**, *35*, 210–211.

(30) Roothaan, C. C. J. "New Developments in Molecular Orbital Theory". Reviews of Modern Physics **1951**, *23*, 69–89.

(31) Hall, G. G.; Lennard-Jones, J. E. "The Molecular Orbital Theory of Chemical Valency VIII. A Method of Calculating Ionization Potentials". Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **1997**, *205*, 541–552.

(32) Davidson, E. R. "The Iterative Calculation of a Few of the Lowest Eigenvalues and Corresponding Eigenvectors of Large Real-Symmetric Matrices". Journal of Computational Physics **1975**, *17*, 87–94.

(33) Lanczos, C. "An Iteration Method for the Solution of the Eigenvalue Problem of Linear Differential and Integral Operators". Journal of Research of the National Bureau of Standards **1950**, *45*, 255.

(34) Arnoldi, W. E. "The Principle of Minimized Iterations in the Solution of the Matrix Eigenvalue Problem". Quarterly of Applied Mathematics **1951**, *9*, 17–29.

(35) Eade, R. H. A.; Robb, M. A. "Direct Minimization in Mc Scf Theory. the Quasi-Newton Method". Chemical Physics Letters **1981**, *83*, 362–368.

(36) Roos, B. "A New Method for Large-Scale Cl Calculations". Chemical Physics Letters **1972**, *15*, 153–159.

(37) Olsen, J. "The CASSCF Method: A Perspective and Commentary". International Journal of Quantum Chemistry **2011**, *111*, 3267–3272.

(38) Roos, B. O.; Taylor, P. R.; Sigbahn, P. E. M. "A Complete Active Space SCF Method (CASSCF) Using a Density Matrix Formulated Super-CI Approach". Chemical Physics **1980**, *48*, 157–173.

(39) Siegbahn, P.; Heiberg, A.; Roos, B.; Levy, B. "A Comparison of the Super-CI and the Newton-Raphson Scheme in the Complete Active Space SCF Method". Physica Scripta **1980**, *21*, 323.

(40) Siegbahn, P. E. M.; Almlöf, J.; Heiberg, A.; Roos, B. O. "The Complete Active Space SCF (CASSCF) Method in a Newton–Raphson Formulation with Application to the HNO Molecule". The Journal of Chemical Physics **1981**, *74*, 2384–2396.

(41) Rayleigh, J. W. S. B.; Lindsay, R. B., *The Theory of Sound*; Courier Corporation: 1945; 548 pp.

(42) Schrödinger, E. "Quantisierung Als Eigenwertproblem". Annalen der Physik **1926**, *385*, 437–490.

(43) Møller, Chr.; Plesset, M. S. "Note on an Approximation Treatment for Many-Electron Systems". Physical Review **1934**, *46*, 618–622.

(44) Čížek, J. "On the Correlation Problem in Atomic and Molecular Systems. Calculation of Wavefunction Components in Ursell-Type Expansion Using Quantum-Field Theoretical Methods". The Journal of Chemical Physics **1966**, *45*, 4256–4266.

(45) Čížek, J.; Paldus, J. "Correlation Problems in Atomic and Molecular Systems III. Rederivation of the Coupled-Pair Many-Electron Theory Using the Traditional Quantum Chemical Methodst". International Journal of Quantum Chemistry **1971**, *5*, 359–379.

(46) Paldus, J.; čížek, J., *Time-Independent Diagrammatic Approach to Perturbation Theory of Fermion Systems* In *Advances in Quantum Chemistry*, Löwdin, P.-O., Ed.; Academic Press: 1975; Vol. 9, pp 105–197.

(47) Shavitt, I.; Bartlett, R. J., *Many-Body Methods in Chemistry and Physics: MBPT and Coupled-Cluster Theory*; Cambridge Molecular Science; Cambridge University Press: Cambridge, 2009.

(48) Aoto, Y. A.; Köhn, A. "Internally Contracted Multireference Coupled-Cluster Theory in a Multistate Framework". The Journal of Chemical Physics **2016**, *144*, 074103.

(49) Evangelista, F. A. "Perspective: Multireference Coupled Cluster Theories of Dynamical Electron Correlation". The Journal of Chemical Physics **2018**, *149*, 030901.

(50) Hanauer, M.; Köhn, A. "Pilot Applications of Internally Contracted Multireference Coupled Cluster Theory, and How to Choose the Cluster Operator Properly". The Journal of Chemical Physics **2011**, *134*, 204111.

(51) Jankowski, K.; Paldus, J.; Grabowski, I.; Kowalski, K. "Applicability of Valence-universal Multireference Coupled-cluster Theories to Quasidegenerate Electronic States. I. Models Involving at Most Two-body Amplitudes". The Journal of Chemical Physics **1992**, *97*, 7600–7612.

(52)   Jeziorski, B.; Monkhorst, H. J. "Coupled-Cluster Method for Multideterminantal Reference States". Physical Review A **1981**, *24*, 1668–1681.

(53)   Köhn, A.; Black, J. A.; Aoto, Y. A.; Hanauer, M. "Improved and Simplified Orthogonalisation Scheme and Connected Triples Correction within the Internally Contracted Multireference Coupled-Cluster Method". Molecular Physics **2020**, *118*, e1743889.

(54)   Kats, D.; Manby, F. R. "Communication: The Distinguishable Cluster Approximation". The Journal of Chemical Physics **2013**, *139*, 021102.

(55)   Kats, D.; Kreplin, D.; Werner, H.-J.; Manby, F. R. "Accurate Thermochemistry from Explicitly Correlated Distinguishable Cluster Approximation". The Journal of Chemical Physics **2015**, *142*, 064111.

(56)   Rishi, V.; Perera, A.; Nooijen, M.; Bartlett, R. J. "Excited States from Modified Coupled Cluster Methods: Are They Any Better than EOM CCSD?" The Journal of Chemical Physics **2017**, *146*, 144104.

(57)   Vitale, E.; Alavi, A.; Kats, D. "FCIQMC-Tailored Distinguishable Cluster Approach". Journal of Chemical Theory and Computation **2020**, *16*, 5621–5634.

(58)   Raghavachari, K.; Trucks, G. W.; Pople, J. A.; Head-Gordon, M. "A Fifth-Order Perturbation Comparison of Electron Correlation Theories". Chemical Physics Letters **1989**, *157*, 479–483.

(59)   Woon, D. E.; Dunning Jr., T. H. "Gaussian Basis Sets for Use in Correlated Molecular Calculations. III. The Atoms Aluminum through Argon". The Journal of Chemical Physics **1993**, *98*, 1358–1371.

(60)   Woon, D. E.; Dunning Jr., T. H. "Gaussian Basis Sets for Use in Correlated Molecular Calculations. IV. Calculation of Static Electrical Response Properties". The Journal of Chemical Physics **1994**, *100*, 2975–2988.

(61)   Wilson, A. K.; van Mourik, T.; Dunning, T. H. "Gaussian Basis Sets for Use in Correlated Molecular Calculations. VI. Sextuple Zeta Correlation Consistent Basis Sets for Boron through Neon". Journal of Molecular Structure: THEOCHEM **1996**, *388*, 339–349.

(62)   Klopper, W.; Manby, F. R.; Ten-No, S.; Valeev, E. F. "R12 Methods in Explicitly Correlated Molecular Electronic Structure Theory". https://doi.org/10.1080/01442350600799921 **2007**, *25*, 427–468.

(63)   Klopper, W.; Schütz, M.; Lüthi, H. P.; Leutwyler, S. "An Ab Initio Derived Torsional Potential Energy Surface for (H2O)3. II. Benchmark Studies and Interaction Energies". The Journal of Chemical Physics **1995**, *103*, 1085–1098.

(64) Grüneis, A.; Hirata, S.; Ohnishi, Y.-y.; Ten-no, S. "Perspective: Explicitly Correlated Electronic Structure Theory for Complex Systems". The Journal of Chemical Physics **2017**, *146*, 080901.

(65) Hättig, C.; Klopper, W.; Köhn, A.; Tew, D. P. "Explicitly Correlated Electrons in Molecules". Chemical Reviews **2011**, *112*, 4–74.

(66) Shiozaki, T.; Werner, H.-J. "Multireference Explicitly Correlated F12 Theories". Molecular Physics **2013**, *111*, 607–630.

(67) Kato, T. "On the Eigenfunctions of Many-Particle Systems in Quantum Mechanics". Communications on Pure and Applied Mathematics **1957**, *10*, 151–177.

(68) Pack, R. T.; Brown, W. B. "Cusp Conditions for Molecular Wavefunctions". The Journal of Chemical Physics **1966**, *45*, 556–559.

(69) Kurokawa, Y. I.; Nakashima, H.; Nakatsuji, H., *Chapter Two - General Coalescence Conditions for the Exact Wave Functions: Higher-Order Relations for Coulombic and Non-Coulombic Systems* In *Advances in Quantum Chemistry*, Hoggan, P. E., Ozdogan, T., Eds.; Electron Correlation in Molecules – Ab Initio Beyond Gaussian Quantum Chemistry, Vol. 73; Academic Press: 2016, pp 59–79.

(70) Roothaan, C. C. J.; Sachs, L. M.; Weiss, A. W. "Analytical Self-Consistent Field Functions for the Atomic Configurations $1s^{2}$, $1s^{2}2s$, and $1s^{2}2s^{2}$". Reviews of Modern Physics **1960**, *32*, 186–194.

(71) Watson, R. E. "Approximate Wave Functions for Atomic Be". Physical Review **1960**, *119*, 170–177.

(72) Weiss, A. W. "Configuration Interaction in Simple Atomic Systems". Physical Review **1961**, *122*, 1826–1836.

(73) Schwartz, C. "Ground State of the Helium Atom". Physical Review **1962**, *128*, 1146–1148.

(74) Coulson, C. A.; Neilson, A. H. "Electron Correlation in the Ground State of Helium". Proceedings of the Physical Society **1961**, *78*, 831.

(75) Gilbert, T. L. "Interpretation of the Rapid Convergence of Correlated Wave Functions". Reviews of Modern Physics **1963**, *35*, 491–494.

(76) Prendergast, D.; Nolan, M.; Filippi, C.; Fahy, S.; Greer, J. C. "Impact of Electron-Electron Cusp on Configuration Interaction Energies". Journal of Chemical Physics **2001**, *115*, 1626–1634.

(77) Kong, L.; Bischoff, F. A.; Valeev, E. F. "Explicitly Correlated R12/F12 Methods for Electronic Structure". Chemical Reviews **2011**, *112*, 75–107.

(78) Grynberg, G.; Aspect, A.; Fabre, C., *Introduction to Quantum Optics: From the Semi-classical Approach to Quantized Light*; Cambridge University Press: Cambridge, 2010.

(79)   Slater, J. C. "Central Fields and Rydberg Formulas in Wave Mechanics". Physical Review **1928**, *31*, 333–343.

(80)   Slater, J. C. "The Normal State of Helium". Physical Review **1928**, *32*, 349–360.

(81)   Hylleraas, E. A. "Über den Grundzustand des Heliumatoms". Zeitschrift für Physik **1928**, *48*, 469–494.

(82)   Hylleraas, E. A. "Neue Berechnung Der Energie Des Heliums Im Grundzustande, Sowie Des Tiefsten Terms von Ortho-Helium". Zeitschrift für Physik 1929 54:5 **1929**, *54*, 347–366.

(83)   Largo-Cabrerizo, A.; Clementi, E. "The Hylleraas-CI Method in Molecular Calculations: Two-electron Integrals". Journal of Computational Chemistry **1987**, *8*, 1191–1198.

(84)   James, H. M.; Coolidge, A. S. "The Ground State of the Hydrogen Molecule". The Journal of Chemical Physics **1933**, *1*, 825–835.

(85)   Kołos, W.; Wolniewicz, L. "Accurate Adiabatic Treatment of the Ground State of the Hydrogen Molecule". The Journal of Chemical Physics **1964**, *41*, 3663–3673.

(86)   Perkins, J. F. "Atomic Integrals Containing R23λ R31μ R12v". The Journal of Chemical Physics **1968**, *48*, 1985–1988.

(87)   Perkins, J. F. "Atomic Integrals Containing Rijp Correlation Factors with Unlinked Indices". The Journal of Chemical Physics **1969**, *50*, 2819–2823.

(88)   Sims, J. S.; Hagstrom, S. "Combined Configuration-Interaction—Hylleraas-Type Wave-Function Study of the Ground State of the Beryllium Atom". Physical Review A **1971**, *4*, 908–916.

(89)   Sims, J. S.; Hagstrom, S. A. "One-Center Rij Integrals Over Slater-Type Orbitals". The Journal of Chemical Physics **1971**, *55*, 4699–4710.

(90)   Clary, D. C.; Handy, N. C. "Hylleraas-Type Wavefunction for Lithium Hydride". Chemical Physics Letters **1977**, *51*, 483–486.

(91)   Clary, D. C.; Handy, N. C. "CI-Hylleraas Variational Calculation on the Ground State of the Neon Atom". Physical Review A **1976**, *14*, 1607–1613.

(92)   Boys, S. F. "The Integral Formulae for the Variational Solution of the Molecular Many-Electron Wave Equation in Terms of Gaussian Functions with Direct Electronic Correlation". Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **1960**, *258*, 402–411.

(93)   Singer, K. "The Use of Gaussian (Exponential Quadratic) Wave Functions in Molecular Problems - I. General Formulae for the Evaluation of Integrals". Proceedings of the Royal Society of London. Series A. Mathematical and Physical Sciences **1960**, *258*, 412–420.

(94) Mitroy, J.; Bubin, S.; Horiuchi, W.; Suzuki, Y.; Adamowicz, L.; Cencek, W.; Szalewicz, K.; Komasa, J.; Blume, D.; Varga, K. "Theory and Application of Explicitly Correlated Gaussians". Reviews of Modern Physics **2013**, *85*, 693–749.

(95) Lester Jr., W. A.; Krauss, M. "Gaussian Correlation Functions: Two-Electron Systems". The Journal of Chemical Physics **1964**, *41*, 1407–1413.

(96) Bukowski, R.; Jeziorski, B.; Szalewicz, K. "New Effective Strategy of Generating Gaussian-type Geminal Basis Sets for Correlation Energy Calculations". The Journal of Chemical Physics **1994**, *100*, 1366–1374.

(97) Persson, B. J.; Taylor, P. R. "Accurate Quantum-chemical Calculations: The Use of Gaussian-type Geminal Functions in the Treatment of Electron Correlation". The Journal of Chemical Physics **1996**, *105*, 5915–5926.

(98) Pan, K.-C.; King, H. F. "Gaussian Geminals for Electron Pair Correlation". The Journal of Chemical Physics **1970**, *53*, 4397–4399.

(99) Pan, K.-C.; King, H. F. "Electron Correlation in Closed Shell Systems. I. Perturbation Theory Using Gaussian-Type Geminals". The Journal of Chemical Physics **1972**, *56*, 4667–4688.

(100) Szalewicz, K.; Jeziorski, B.; Monkhorst, H. J.; Zabolitzky, J. G. "A New Functional for Variational Calculation of Atomic and Molecular Second-Order Correlation Energies". Chemical Physics Letters **1982**, *91*, 169–172.

(101) Szalewicz, K.; Jeziorski, B.; Monkhorst, H. J.; Zabolitzky, J. G. "Atomic and Molecular Correlation Energies with Explicitly Correlated Gaussian Geminals. I. Second-order Perturbation Treatment for He, Be, H2, and LiH". The Journal of Chemical Physics **1983**, *78*, 1420–1430.

(102) Wenzel, K. B.; Zabolitzky, J. G.; Szalewicz, K.; Jeziorski, B.; Monkhorst, H. J. "Atomic and Molecular Correlation Energies with Explicitly Correlated Gaussian Geminals. V. Cartesian Gaussian Geminals and the Neon Atom". The Journal of Chemical Physics **1986**, *85*, 3964–3974.

(103) Szalewicz, K.; Zabolitzky, J. G.; Jeziorski, B.; Monkhorst, H. J. "Atomic and Molecular Correlation Energies with Explicitly Correlated Gaussian Geminals. IV. A Simplified Treatment of Strong Orthogonality in MBPT and Coupled Cluster Calculations". The Journal of Chemical Physics **1984**, *81*, 2723–2731.

(104) Tew, D. P.; Klopper, W.; Manby, F. R. "The Weak Orthogonality Functional in Explicitly Correlated Pair Theories". The Journal of Chemical Physics **2007**, *127*, 174105.

(105) Korobov, V. I. "Coulomb Three-Body Bound-State Problem: Variational Calculations of Nonrelativistic Energies". Physical Review A **2000**, *61*, 064503.

(106) Varga, K.; Suzuki, Y. "Precise Solution of Few-Body Problems with the Stochastic Variational Method on a Correlated Gaussian Basis". Physical Review C **1995**, *52*, 2885–2905.

(107) Bubin, S.; Varga, K. "Ground-State Energy and Relativistic Corrections for Positronium Hydride". Physical Review A **2011**, *84*, 012509.

(108) Stanke, M.; Bubin, S.; Molski, M.; Adamowicz, L. "Non-Born-Oppenheimer Calculations of the Lowest Vibrational Energy of HD Including Relativistic Corrections". Physical Review A **2009**, *79*, 032507.

(109) Kutzelnigg, W. "R 12-Dependent Terms in the Wave Function as Closed Sums of Partial Wave Amplitudes for Large l". Theoretica chimica acta 1985 68:6 **1985**, *68*, 445–469.

(110) Ten-no, S. "Initiation of Explicitly Correlated Slater-type Geminal Theory". Chemical Physics Letters **2004**, *398*, 56–61.

(111) Werner, H.-J.; Adler, T. B.; Manby, F. R. "General Orbital Invariant MP2-F12 Theory". The Journal of Chemical Physics **2007**, *126*, 164102.

(112) Tew, D. P.; Kats, D. "Relaxing Constrained Amplitudes: Improved F12 Treatments of Orbital Optimization and Core–Valence Correlation Energies". Journal of Chemical Theory and Computation **2018**, *14*, 5435–5440.

(113) Johnson, C. M.; Hirata, S.; Ten-no, S. "Explicit Correlation Factors". Chemical Physics Letters **2017**, *683*, 247–252.

(114) Hylleraas, E. A. "Über den Grundterm der Zweielektronenprobleme von H-, He, Li+, Be++ usw." Zeitschrift für Physik **1930**, *65*, 209–225.

(115) Bethe, H. A.; Salpeter, E. E., *Quantum Mechanics of One- and Two-Electron Atoms*; Springer: Berlin, Heidelberg, 1957.

(116) Klopper, W.; Kutzelnigg, W. "Møller-Plesset Calculations Taking Care of the Correlation CUSP". Chemical Physics Letters **1987**, *134*, 17–22.

(117) Kutzelnigg, W.; Klopper, W. "Wave Functions with Terms Linear in the Interelectronic Coordinates to Take Care of the Correlation Cusp. I. General Theory". The Journal of Chemical Physics **1991**, *94*, 1985–2001.

(118) Klopper, W. "Orbital-Invariant Formulation of the MP2-R12 Method". Chemical Physics Letters **1991**, *186*, 583–585.

(119) Tew, D. P.; Klopper, W. "Open-Shell Explicitly Correlated F12 Methods". Molecular Physics **2010**, *108*, 315–325.

(120) Bokhan, D.; Ten-no, S.; Noga, J. "Implementation of the CCSD(T)-F12 Method Using Cusp Conditions". Physical Chemistry Chemical Physics **2008**, *10*, 3320–3326.

(121) Ten-no, S. "New Implementation of Second-Order Møller-Plesset Perturbation Theory with an Analytic Slater-type Geminal". The Journal of Chemical Physics **2007**, *126*, 014108.

(122) Tew, D. P.; Klopper, W. "A Comparison of Linear and Nonlinear Correlation Factors for Basis Set Limit Møller-Plesset Second Order Binding Energies and Structures of He2, Be2, and Ne2". The Journal of Chemical Physics **2006**, *125*, 094302.

(123) Klopper, W.; Samson, C. C. M. "Explicitly Correlated Second-Order Møller–Plesset Methods with Auxiliary Basis Sets". The Journal of Chemical Physics **2002**, *116*, 6397–6410.

(124) Adler, T. B.; Knizia, G.; Werner, H.-J. "A Simple and Efficient CCSD(T)-F12 Approximation". The Journal of Chemical Physics **2007**, *127*, 221106.

(125) Kutzelnigg, W. "R12-Dependent Terms in the Wave Function as Closed Sums of Partial Wave Amplitudes for Large l". Theoretica chimica acta **1985**, *68*, 445–469.

(126) Klopper, W.; Kutzelnigg, W. "MP2-R12 Calculations on the Relative Stability of Carbocations".

(127) Termath, V.; Klopper, W.; Kutzelnigg, W. "Wave Functions with Terms Linear in the Interelectronic Coordinates to Take Care of the Correlation Cusp. II. Second-order Mo/Ller–Plesset (MP2-R12) Calculations on Closed-shell Atoms". The Journal of Chemical Physics **1991**, *94*, 2002–2019.

(128) Wind, P.; Klopper, W.; Helgaker, T. "Second-Order Møller–Plesset Perturbation Theory with Terms Linear in the Interelectronic Coordinates and Exact Evaluation of Three-Electron Integrals". Theoretical Chemistry Accounts **2002**, *107*, 173–179.

(129) Valeev, E. F. "Improving on the Resolution of the Identity in Linear R12 Ab Initio Theories". Chemical Physics Letters **2004**, *395*, 190–195.

(130) Klopper, W.; Kutzelnigg, W. "Wave Functions with Terms Linear in the Interelectronic Coordinates to Take Care of the Correlation Cusp. III. Second-order Mo/Ller–Plesset (MP2-R12) Calculations on Molecules of First Row Atoms". The Journal of Chemical Physics **1991**, *94*, 2020–2030.

(131) Manby, F. R. "Density Fitting in Second-Order Linear-R12 Møller–Plesset Perturbation Theory". The Journal of Chemical Physics **2003**, *119*, 4607–4613.

(132) Ten-no, S.; Manby, F. R. "Density Fitting for the Decomposition of Three-Electron Integrals in Explicitly Correlated Electronic Structure Theory". The Journal of Chemical Physics **2003**, *119*, 5358–5363.

(133) Kedžuch, S.; Milko, M.; Noga, J. "Alternative Formulation of the Matrix Elements in MP2-R12 Theory". International Journal of Quantum Chemistry **2005**, *105*, 929–936.

(134)   Shiozaki, T.; Kamiya, M.; Hirata, S.; Valeev, E. F. "Explicitly Correlated Coupled-Cluster Singles and Doubles Method Based on Complete Diagrammatic Equations". The Journal of Chemical Physics **2008**, *129*, DOI: `10.1063/1.2967181`.

(135)   Köhn, A.; Richings, G. W.; Tew, D. P. "Implementation of the Full Explicitly Correlated Coupled-Cluster Singles and Doubles Model CCSD-F12 with Optimally Reduced Auxiliary Basis Dependence". The Journal of Chemical Physics **2008**, *129*, DOI: `10.1063/1.3028546`.

(136)   Shiozaki, T.; Kamiya, M.; Hirata, S.; Valeev, E. F. "Higher-Order Explicitly Correlated Coupled-Cluster Methods". The Journal of Chemical Physics **2009**, *130*, 054101.

(137)   Werner, H.-J.; Manby, F. R. "Explicitly Correlated Second-Order Perturbation Theory Using Density Fitting and Local Approximations". The Journal of Chemical Physics **2006**, *124*, 054114.

(138)   Manby, F. R.; Werner, H.-J.; Adler, T. B.; May, A. J. "Explicitly Correlated Local Second-Order Perturbation Theory with a Frozen Geminal Correlation Factor". The Journal of Chemical Physics **2006**, *124*, 094103.

(139)   Werner, H.-J. "Eliminating the Domain Error in Local Explicitly Correlated Second-Order Møller–Plesset Perturbation Theory". The Journal of Chemical Physics **2008**, *129*, DOI: `10.1063/1.2982419`.

(140)   Peterson, K. A.; Adler, T. B.; Werner, H.-J. "Systematically Convergent Basis Sets for Explicitly Correlated Wavefunctions: The Atoms H, He, B–Ne, and Al–Ar". The Journal of Chemical Physics **2008**, *128*, DOI: `10.1063/1.2831537`.

(141)   Yousaf, K. E.; Peterson, K. A. "Optimized Auxiliary Basis Sets for Explicitly Correlated Methods". The Journal of Chemical Physics **2008**, *129*, DOI: `10.1063/1.3009271`.

(142)   Yousaf, K. E.; Peterson, K. A. "Optimized Complementary Auxiliary Basis Sets for Explicitly Correlated Methods: Aug-Cc-pVnZ Orbital Basis Sets". Chemical Physics Letters **2009**, *476*, 303–307.

(143)   Ten-no, S. "Explicitly Correlated Second Order Perturbation Theory: Introduction of a Rational Generator and Numerical Quadratures". The Journal of Chemical Physics **2004**, *121*, 117–129.

(144)   Shiozaki, T. "Evaluation of Slater-type Geminal Integrals Using Tailored Gaussian Quadrature". Chemical Physics Letters **2009**, *479*, 160–164.

(145)   May, A. J.; Manby, F. R. "An Explicitly Correlated Second Order Møller-Plesset Theory Using a Frozen Gaussian Geminal". The Journal of Chemical Physics **2004**, *121*, 4479–4485.

(146) Adler, T. B.; Werner, H.-J.; Manby, F. R. "Local Explicitly Correlated Second-Order Perturbation Theory for the Accurate Treatment of Large Molecules". The Journal of Chemical Physics **2009**, *130*, DOI: `10.1063/1.3040174`.

(147) Klopper, W. "A Hybrid Scheme for the Resolution-of-the-Identity Approximation in Second-Order Møller–Plesset Linear-R12 Perturbation Theory". The Journal of Chemical Physics **2004**, *120*, 10890–10895.

(148) Noga, J.; Kutzelnigg, W.; Klopper, W. "CC-R12, a Correlation Cusp Corrected Coupled-Cluster Method with a Pilot Application to the Be2 Potential Curve". Chemical Physics Letters **1992**, *199*, 497–504.

(149) Noga, J.; Kutzelnigg, W. "Coupled Cluster Theory That Takes Care of the Correlation Cusp by Inclusion of Linear Terms in the Interelectronic Coordinates". The Journal of Chemical Physics **1994**, *101*, 7738–7762.

(150) Gdanitz, R. J. "A Formulation of Multiple-Reference CI with Terms Linear in the Interelectronic Distances". Chemical Physics Letters **1993**, *210*, 253–260.

(151) Gdanitz, R. J. "Accurately Solving the Electronic Schrödinger Equation of Atoms and Molecules Using Explicitly Correlated (R12-)MR-CI: The Ground State Potential Energy Curve of N2". Chemical Physics Letters **1998**, *283*, 253–261.

(152) Fliegl, H.; Hättig, C.; Klopper, W. "Coupled-Cluster Response Theory with Linear-R12 Corrections: The CC2-R12 Model for Excitation Energies". The Journal of Chemical Physics **2006**, *124*, 044112.

(153) Neiss, C.; Hättig, C.; Klopper, W. "Extensions of R12 Corrections to CC2-R12 for Excited States". The Journal of Chemical Physics **2006**, *125*, 064111.

(154) Köhn, A. "A Modified Ansatz for Explicitly Correlated Coupled-Cluster Wave Functions That Is Suitable for Response Theory". The Journal of Chemical Physics **2009**, *130*, 104104.

(155) Bokhan, D.; Ten-no, S. "Communications: Explicitly Correlated Equation-of-Motion Coupled Cluster Method for Ionized States". The Journal of Chemical Physics **2010**, *132*, 021101.

(156) Höfener, S.; Schieschke, N.; Klopper, W.; Köhn, A. "The Extended Explicitly-Correlated Second-Order Approximate Coupled-Cluster Singles and Doubles Ansatz Suitable for Response Theory". The Journal of Chemical Physics **2019**, *150*, 184110.

(157) Kedžuch, S.; Demel, O.; Pittner, J.; Ten-no, S.; Noga, J. "Multireference F12 Coupled Cluster Theory: The Brillouin-Wigner Approach with Single and Double Excitations". Chemical Physics Letters **2011**, *511*, 418–423.

(158)   Neese, F.; Wennmohs, F.; Hansen, A. "Efficient and Accurate Local Approximations to Coupled-Electron Pair Approaches: An Attempt to Revive the Pair Natural Orbital Method". The Journal of Chemical Physics **2009**, *130*, 114108.

(159)   Werner, H.-J.; Köppl, C.; Ma, Q.; Schwilk, M., *Explicitly Correlated Local Electron Correlation Methods* In *Fragmentation*; John Wiley & Sons, Ltd: 2017; Chapter 1, pp 1–79.

(160)   Varganov, S. A.; Martínez, T. J. "Variational Geminal-Augmented Multireference Self-Consistent Field Theory: Two-electron Systems". The Journal of Chemical Physics **2010**, *132*, 054103.

(161)   Ten-no, S. "A Simple F12 Geminal Correction in Multi-Reference Perturbation Theory". Chemical Physics Letters **2007**, *447*, 175–179.

(162)   Shiozaki, T.; Werner, H.-J. "Communication: Second-order Multireference Perturbation Theory with Explicit Correlation: CASPT2-F12". The Journal of Chemical Physics **2010**, *133*, 141103.

(163)   Sharma, S.; Yanai, T.; Booth, G. H.; Umrigar, C. J.; Chan, G. K.-L. "Spectroscopic Accuracy Directly from Quantum Chemistry: Application to Ground and Excited States of Beryllium Dimer". The Journal of Chemical Physics **2014**, *140*, 104112.

(164)   Knizia, G.; Adler, T. B.; Werner, H.-J. "Simplified CCSD(T)-F12 Methods: Theory and Benchmarks". The Journal of Chemical Physics **2009**, *130*, 054104.

(165)   Flores, J. R.; Gdanitz, R. J. "Accurately Solving the Electronic Schrödinger Equation of Small Atoms and Molecules Using Explicitly Correlated (R12-)MR-CI. VIII. Valence Excited States of Methylene (CH2)". The Journal of Chemical Physics **2005**, *123*, 144316.

(166)   Shiozaki, T.; Werner, H.-J. "Explicitly Correlated Multireference Configuration Interaction with Multiple Reference Functions: Avoided Crossings and Conical Intersections". The Journal of Chemical Physics **2011**, *134*, 184104.

(167)   Hirschfelder, J. O. "Removal of Electron—Electron Poles from Many-Electron Hamiltonians". The Journal of Chemical Physics **1963**, *39*, 3145–3146.

(168)   Jankowski, K. "A Perturbation Treatment of the Many-Electron Problem". Acta Physica Polonica **1967**, *XXXII*, 421.

(169)   Jankowski, K. "A Perturbation Treatment of the Many-Electron Problem II. Third Order Solutions for the Ground State of Heliumlike Systems". Acta Physica Polonica, A **1970**, *37*, 669.

(170)   Boys, S. F.; Handy, N. C. "A Calculation for the Energies and Wavefunctions for States of Neon with Full Electronic Correlation Accuracy". Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **1969**, *310*, 63–78.

(171) Boys, S. F.; Handy, N. C. "A Condition to Remove the Indeterminacy in Inter-electronic Correlation Functions". Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **1969**, *309*, 209–220.

(172) Boys, S. F.; Handy, N. C. "The Determination of Energies and Wavefunctions with Full Electronic Correlation". Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **1969**, *310*, 43–61.

(173) Boys, S. F.; Handy, N. C. "A First Solution, for LiH, of a Molecular Transcorrelated Wave Equation by Means of Restricted Numerical Integration". Proceedings of the Royal Society of London. A. Mathematical and Physical Sciences **1969**, *311*, 309–329.

(174) Jastrow, R. "Many-Body Problem with Strong Forces". Physical Review **1955**, *98*, 1479.

(175) Cohen, A. J.; Luo, H.; Guther, K.; Dobrautz, W.; Tew, D. P.; Alavi, A. "Similarity Transformation of the Electronic Schrödinger Equation via Jastrow Factorization". Journal of Chemical Physics **2019**, *151*, 061101.

(176) Handy, N. C. "Energies and Expectation Values for Be by the Transcorrelated Method". The Journal of Chemical Physics **1969**, *51*, 3205–3212.

(177) Nooijen, M.; Bartlett, R. J. "Elimination of Coulombic Infinities through Transformation of the Hamiltonian". The Journal of Chemical Physics **1998**, *109*, 8232–8240.

(178) Ten-No, S. "A Feasible Transcorrelated Method for Treating Electronic Cusps Using a Frozen Gaussian Geminal". Chemical Physics Letters **2000**, *330*, 169–174.

(179) Ten-no, S. "Three-Electron Integral Evaluation in the Transcorrelated Method Using a Frozen Gaussian Geminal". Chemical Physics Letters **2000**, *330*, 175–179.

(180) Zweistra, H. J. A.; Samson, C. C. M.; Klopper, W. "Similarity-Transformed Hamiltonians by Means of Gaussian-Damped Interelectronic Distances". Collection of Czechoslovak Chemical Communications **2003**, *68*, 374–386.

(181) Luo, H.; Hackbusch, W.; Flad, H. J. "Quantum Monte Carlo Study of the Transcorrelated Method for Correlation Factors". Molecular Physics **2010**, *108*, 425–431.

(182) Luo, H. "Variational Transcorrelated Method". The Journal of Chemical Physics **2010**, *133*, 154109.

(183) Luo, H. "Complete Optimisation of Multi-Configuration Jastrow Wave Functions by Variational Transcorrelated Method". The Journal of Chemical Physics **2011**, *135*, 024109.

(184) Luo, H.; Hackbusch, W.; Flad, H.-J. "Quantum Monte Carlo Study of the Transcorrelated Method for Correlation Factors". Molecular Physics **2010**, *108*, 425–431.

(185)   Ammar, A.; Scemama, A.; Loos, P.-F.; Giner, E. "Compactification of Determinant Expansions via Transcorrelation". The Journal of Chemical Physics **2024**, *161*, 084104.

(186)   Ammar, A.; Scemama, A.; Giner, E. "Extension of Selected Configuration Interaction for Transcorrelated Methods". The Journal of Chemical Physics **2022**, *157*, 134107.

(187)   Ammar, A.; Scemama, A.; Giner, E. "Transcorrelated Selected Configuration Interaction in a Bi-Orthonormal Basis and with a Cheap Three-Body Correlation Factor". The Journal of Chemical Physics **2023**, *159*, 114121.

(188)   Baiardi, A.; Lesiuk, M.; Reiher, M. "Explicitly Correlated Electronic Structure Calculations with Transcorrelated Matrix Product Operators". Journal of Chemical Theory and Computation **2022**, *18*, 4203–4217.

(189)   Baiardi, A.; Reiher, M. "Transcorrelated Density Matrix Renormalization Group". The Journal of Chemical Physics **2020**, *153*, 164115.

(190)   Christlmaier, E. M. C.; Schraivogel, T.; López Ríos, P.; Alavi, A.; Kats, D. "xTC: An Efficient Treatment of Three-Body Interactions in Transcorrelated Methods". The Journal of Chemical Physics **2023**, *159*, 014113.

(191)   Dobrautz, W.; Luo, H.; Alavi, A. "Compact Numerical Solutions to the Two-Dimensional Repulsive Hubbard Model Obtained via Nonunitary Similarity Transformations". Physical Review B **2019**, *99*, 075119.

(192)   Dobrautz, W.; Cohen, A. J.; Alavi, A.; Giner, E. "Performance of a One-Parameter Correlation Factor for Transcorrelation: Study on a Series of Second Row Atomic and Molecular Systems". Journal of Chemical Physics **2022**, *156*, 234108.

(193)   Dobrautz, W.; Sokolov, I. O.; Liao, K.; López Ríos, P.; Rahm, M.; Alavi, A.; Tavernelli, I. "Toward Real Chemical Accuracy on Current Quantum Hardware Through the Transcorrelated Method". Journal of Chemical Theory and Computation **2024**, *20*, 4146–4160.

(194)   Giner, E. "A New Form of Transcorrelated Hamiltonian Inspired by Range-Separated DFT". Journal of Chemical Physics **2021**, *154*, DOI: `10.1063/5.0044683`.

(195)   Guther, K.; Cohen, A. J.; Luo, H.; Alavi, A. "Binding Curve of the Beryllium Dimer Using Similarity-Transformed FCIQMC: Spectroscopic Accuracy with Triple-Zeta Basis Sets". The Journal of Chemical Physics **2021**, *155*, 011102.

(196)   Hino, O.; Tanimura, Y.; Ten-no, S. "Application of the Transcorrelated Hamiltonian to the Linearized Coupled Cluster Singles and Doubles Model". Chemical Physics Letters **2002**, *353*, 317–323.

(197) Imamura, Y.; Scuseria, G. E. "A New Correlation Functional Based on a Transcorrelated Hamiltonian". The Journal of Chemical Physics **2003**, *118*, 2464–2469.

(198) Jeszenszki, P.; Luo, H.; Alavi, A.; Brand, J. "Accelerating the Convergence of Exact Diagonalization with the Transcorrelated Method: Quantum Gas in One Dimension with Contact Interactions". Physical Review A **2018**, *98*, 053627.

(199) Jeszenszki, P.; Ebling, U.; Luo, H.; Alavi, A.; Brand, J. "Eliminating the Wave-Function Singularity for Ultracold Atoms by a Similarity Transformation". Physical Review Research **2020**, *2*, 043270.

(200) Kats, D.; Christlmaier, E. M. C.; Schraivogel, T.; Alavi, A. "Orbital Optimisation in xTC Transcorrelated Methods". Faraday Discussions **2024**, *254*, 382–401.

(201) Kumar, A.; Asthana, A.; Masteran, C.; Valeev, E. F.; Zhang, Y.; Cincio, L.; Tretiak, S.; Dub, P. A. "Quantum Simulation of Molecular Electronic States with a Transcorrelated Hamiltonian: Higher Accuracy with Fewer Qubits". Journal of Chemical Theory and Computation **2022**, *18*, 5312–5324.

(202) Lee, N.; Thom, A. J. W. "Studies on the Transcorrelated Method". Journal of Chemical Theory and Computation **2023**, *19*, 5743–5759.

(203) Liao, K.; Zhai, H.; Christlmaier, E. M. C.; Schraivogel, T.; López Ríos, P.; Kats, D.; Alavi, A. "Density Matrix Renormalization Group for Transcorrelated Hamiltonians: Ground and Excited States in Molecules". Journal of Chemical Theory and Computation **2023**, *19*, 1734–1743.

(204) Liao, K.; Schraivogel, T.; Luo, H.; Kats, D.; Alavi, A. "Towards Efficient and Accurate Ab Initio Solutions to Periodic Systems via Transcorrelation and Coupled Cluster Theory". Physical Review Research **2021**, *3*, 033072.

(205) Luo, H.; Alavi, A. "Combining the Transcorrelated Method with Full Configuration Interaction Quantum Monte Carlo: Application to the Homogeneous Electron Gas". Journal of Chemical Theory and Computation **2018**, *14*, 1403–1411.

(206) Luo, H. "Transcorrelated Calculations of Homogeneous Electron Gases". The Journal of Chemical Physics **2012**, *136*, 224111.

(207) Magnusson, E.; Fitzpatrick, A.; Knecht, S.; Rahm, M.; Dobrautz, W. "Towards Efficient Quantum Computing for Quantum Chemistry: Reducing Circuit Complexity with Transcorrelated and Adaptive Ansatz Techniques". Faraday Discussions **2024**, *254*, 402–428.

(208) Motta, M.; Gujarati, T. P.; Rice, J. E.; Kumar, A.; Masteran, C.; Latone, J. A.; Lee, E.; Valeev, E. F.; Takeshita, T. Y. "Quantum Simulation of Electronic Structure with a Transcorrelated Hamiltonian: Improved Accuracy with a Smaller Footprint on the Quantum Computer". Physical Chemistry Chemical Physics **2020**, *22*, 24270–24281.

(209) Ochi, M.; Sodeyama, K.; Sakuma, R.; Tsuneyuki, S. "Efficient Algorithm of the Transcorrelated Method for Periodic Systems". Journal of Chemical Physics **2012**, *136*, 094108.

(210) Ochi, M.; Yamamoto, Y.; Arita, R.; Tsuneyuki, S. "Iterative Diagonalization of the Non-Hermitian Transcorrelated Hamiltonian Using a Plane-Wave Basis Set: Application to Sp-Electron Systems with Deep Core States". The Journal of Chemical Physics **2016**, *144*, 104109.

(211) Ochi, M.; Tsuneyuki, S. "Optical Absorption Spectra Calculated from a First-Principles Wave Function Theory for Solids: Transcorrelated Method Combined with Configuration Interaction Singles". Journal of Chemical Theory and Computation **2014**, *10*, 4098–4103.

(212) Sakuma, R.; Tsuneyuki, S. "Electronic Structure Calculations of Solids with a Similarity-Transformed Hamiltonian". Journal of the Physical Society of Japan **2006**, *75*, 103705.

(213) Schraivogel, T.; Cohen, A. J.; Alavi, A.; Kats, D. "Transcorrelated Coupled Cluster Methods". The Journal of Chemical Physics **2021**, *155*, 191101.

(214) Schraivogel, T.; Christlmaier, E. M. C.; López Ríos, P.; Alavi, A.; Kats, D. "Transcorrelated Coupled Cluster Methods. II. Molecular Systems". The Journal of Chemical Physics **2023**, *158*, 214106.

(215) Sokolov, I. O.; Dobrautz, W.; Luo, H.; Alavi, A.; Tavernelli, I. "Orders of Magnitude Increased Accuracy for Quantum Many-Body Problems on Quantum Computers via an Exact Transcorrelated Method". Physical Review Research **2023**, *5*, 023174.

(216) Szenes, K.; Moerchen, M.; Fischill, P.; Reiher, M. "Striking the Right Balance of Encoding Electron Correlation in the Hamiltonian and Wavefunction Ansatz". Faraday Discussions **2024**, DOI: `10.1039/D4FD00060A`.

(217) Ten-no, S.; Hino, O. "New Transcorrelated Method Improving the Feasibility of Explicitly Correlated Calculations". International Journal of Molecular Sciences **2002**, *3*, 459–474.

(218) Ten-no, S. L. "Nonunitary Projective Transcorrelation Theory Inspired by the F12 Ansatz". The Journal of Chemical Physics **2023**, *159*, 171103.

(219) Tsuneyuki, S. "Transcorrelated Method: Another Possible Way towards Electronic Structure Calculation of Solids". Progress of Theoretical Physics Supplement **2008**, *176*, 134–142.

(220) Umezawa, N.; Tsuneyuki, S. "Excited Electronic State Calculations by the Transcorrelated Variational Monte Carlo Method: Application to a Helium Atom". The Journal of Chemical Physics **2004**, *121*, 7070–7075.

(221)   Umezawa, N.; Tsuneyuki, S. "Ground-State Correlation Energy for the Homogeneous Electron Gas Calculated by the Transcorrelated Method". Physical Review B **2004**, *69*, 165102.

(222)   Umezawa, N.; Tsuneyuki, S.; Ohno, T.; Shiraishi, K.; Chikyow, T. "A Practical Treatment for the Three-Body Interactions in the Transcorrelated Variational Monte Carlo Method: Application to Atoms from Lithium to Neon". The Journal of Chemical Physics **2005**, *122*, 224101.

(223)   Umezawa, N.; Tsuneyuki, S. "Transcorrelated Method for Electronic Systems Coupled with Variational Monte Carlo Calculation". Journal of Chemical Physics **2003**, *119*, 10015–10031.

(224)   Drummond, N. D.; Towler, M. D.; Needs, R. J. "Jastrow Correlation Factor for Atoms, Molecules, and Solids". DOI: `10.1103/PhysRevB.70.235119`.

(225)   López Ríos, P.; Seth, P.; Drummond, N. D.; Needs, R. J. "Framework for Constructing Generic Jastrow Correlation Factors". PHYSICAL REVIEW E **2012**, *86*, 36703.

(226)   Boyle, P. P. "Options: A Monte Carlo Approach". Journal of Financial Economics **1977**, *4*, 323–338.

(227)   Johnson, W., *Monte Carlo Methods in Finance: Simulation Techniques for Market Modeling*; HiTeX Press: 2024; 454 pp.

(228)   Kroese, D. P.; Brereton, T.; Taimre, T.; Botev, Z. I. "Why the Monte Carlo Method Is so Important Today". WIREs Computational Statistics **2014**, *6*, 386–392.

(229)   Rosenbluth, M. N.; Rosenbluth, A. W. "Monte Carlo Calculation of the Average Extension of Molecular Chains". The Journal of Chemical Physics **1955**, *23*, 356–359.

(230)   Foulkes, W. M. C.; Mitas, L.; Needs, R. J.; Rajagopal, G. "Quantum Monte Carlo Simulations of Solids".

(231)   Thijssen, J., *Computational Physics*; Cambridge University Press: 2007.

(232)   Metropolis, N.; Ulam, S. "The Monte Carlo Method". Journal of the American Statistical Association **1949**, *44*, 335–341.

(233)   Metropolis, N.; Rosenbluth, A. W.; Rosenbluth, M. N.; Teller, A. H.; Teller, E. "Equation of State Calculations by Fast Computing Machines". The Journal of Chemical Physics **1953**, *21*, 1087–1092.

(234)   Leclerc, G.-L. "Essai D'arithmetique Morale".

(235)   *Histoire de l'Académie royale des sciences*, 1735; 696 pp.

(236)   Seneta, E.; Parshall, K. H.; Jongmans, F. "Nineteenth-Century Developments in Geometric Probability: J. J. Sylvester, M. W. Crofton, J.-É. Barbier, and J. Bertrand". Archive for History of Exact Sciences **2001**, *55*, 501–524.

(237)   Allen, M. P.; Tildesley, D. J., *Computer Simulation of Liquids*; Clarendon Press: 1989; 412 pp.

(238)   *Monte Carlo Methods in Statistical Physics*; Binder, K., Ed.; Topics in Current Physics, Vol. 7; Springer: Berlin, Heidelberg, 1986.

(239)   Binder, K.; Heermann, D. W., *Monte Carlo Simulation in Statistical Physics*; Cardona, M., Fulde, P., Von Klitzing, K., Queisser, H.-J., Lotsch, H. K. V., red.; Springer Series in Solid-State Sciences, Vol. 80; Springer: Berlin, Heidelberg, 1988.

(240)   Newman, M. E. J.; Barkema, G. T.; Newman, M. E. J.; Barkema, G. T., *Monte Carlo Methods in Statistical Physics*; Oxford University Press: Oxford, New York, 1999; 490 pp.

(241)   Reiher, W. "Hammersley, J. M., D. C. Handscomb: Monte Carlo Methods. Methuen & Co., London, and John Wiley & Sons, New York, 1964. VII + 178 S., Preis: 25 s". Biometrische Zeitschrift **1966**, *8*, 209–209.

(242)   Bär, H.-J.; Bopp, P. "M. H. Kalos and P. A. Whitlock: Monte Carlo Methods, Volume I: Basics, John Wiley and Sons, New York, Chichester, Brisbane, Toronto and Singapore 1986, Library of Congress QA298.K35 1986. 186 Seiten Mit Einem Index, Preis: £ 29.00". Berichte der Bunsengesellschaft für physikalische Chemie **1988**, *92*, 560–560.

(243)   Bernoulli, J., *Jacobi Bernoulli profess. basil. & utriusque societ. ... Ars conjectandi, opus posthumum : accedit Tractatus de seriebus infinitis, et epistola Gallice scripta de ludo pilæ reticularis*; Basileæ : Impensis Thurnisiorum, Fratrum: 1713; 372 pp.

(244)   James, F. "Monte Carlo Theory and Practice". Reports on Progress in Physics **1980**, *43*, 1145.

(245)   Kahn, H. *Modification of the Monte Carlo*. RAND Corporation, 1959.

(246)   Arouna, B. "Adaptative Monte Carlo Method, A Variance Reduction Technique". **2004**, *10*, 1–24.

(247)   Hastings, W. K. "Monte Carlo Sampling Methods Using Markov Chains and Their Applications". Biometrika **1970**, *57*, 97–109.

(248)   Markov, A. A., *Theory of Algorithms*; Academy of Sciences of the USSR: 1954; 468 pp.

(249)   Flyvbjerg, H.; Petersen, H. G. "Error Estimates on Averages of Correlated Data". The Journal of Chemical Physics **1989**, *91*, 461–466.

(250)   Cuzzocrea, A.; Scemama, A.; Briels, W. J.; Moroni, S.; Filippi, C. "Variational Principles in Quantum Monte Carlo: The Troubled Story of Variance Minimization". Journal of Chemical Theory and Computation **2020**, *16*, 4203–4212.

(251) Snajdr, M.; Rothstein, S. M. "Are Properties Derived from Variance-Optimized Wave Functions Generally More Accurate? Monte Carlo Study of Non-Energy-Related Properties of H2, He, and LiH". The Journal of Chemical Physics **2000**, *112*, 4935–4941.

(252) Kent, P. R. C.; Needs, R. J.; Rajagopal, G. "Monte Carlo Energy and Variance-Minimization Techniques for Optimizing Many-Body Wave Functions". Physical Review B **1999**, *59*, 12344–12351.

(253) Needs, R. J.; Towler, M. D.; Drummond, N. D.; López Ríos, P.; Trail, J. R. "Variational and Diffusion Quantum Monte Carlo Calculations with the CASINO Code". Journal of Chemical Physics **2020**, *152*, 154106.

(254) Fahy, S.; Wang, X. W.; Louie, S. G. "Variational Quantum Monte Carlo Nonlocal Pseudopotential Approach to Solids: Cohesive and Structural Properties of Diamond". Physical Review Letters **1988**, *61*, 1631–1634.

(255) Li, X.-P.; Ceperley, D. M.; Martin, R. M. "Cohesive Energy of Silicon by the Green's-Function Monte Carlo Method". Physical Review B **1991**, *44*, 10929–10932.

(256) Rajagopal, G.; Needs, R. J.; Kenny, S.; Foulkes, W. M. C.; James, A. "Quantum Monte Carlo Calculations for Solids Using Special k Points Methods". Physical Review Letters **1994**, *73*, 1959–1962.

(257) Malatesta, A.; Fahy, S.; Bachelet, G. B. "Variational Quantum Monte Carlo Calculation of the Cohesive Properties of Cubic Boron Nitride". Physical Review B **1997**, *56*, 12201–12210.

(258) Harrison, R. J.; Handy, N. C. "Quantum Monte Carlo Calculations on Be and LiH". Chemical Physics Letters **1985**, *113*, 257–263.

(259) Lester, W. A., *Recent Advances In Quantum Monte Carlo Methods*; World Scientific: 1997; 244 pp.

(260) Umrigar, C. J.; Nightingale, M. P.; Runge, K. J. "A Diffusion Monte Carlo Algorithm with Very Small Time-step Errors". The Journal of Chemical Physics **1993**, *99*, 2865–2890.

(261) Fick, A. "On Liquid Diffusion". Journal of Membrane Science **1995**, *100*, 33–38.

(262) Kosztin, I.; Faber, B.; Schulten, K. "Introduction to the Diffusion Monte Carlo Method". American Journal of Physics **1996**, *64*, 633–644.

(263) Lie, S.; Engel, F., *Theorie der Transformationsgruppen*; American Mathematical Soc.: 1970; 672 pp.

(264) Trotter, H. F. "On the Product of Semi-Groups of Operators". Proceedings of the American Mathematical Society **1959**, *10*, 545–551.

(265)   Suzuki, M. "Generalized Trotter's Formula and Systematic Approximants of Exponential Operators and Inner Derivations with Applications to Many-Body Problems". Communications in Mathematical Physics **1976**, *51*, 183–190.

(266)   Booth, G. H.; Thom, A. J.; Alavi, A. "Fermion Monte Carlo without Fixed Nodes: A Game of Life, Death, and Annihilation in Slater Determinant Space". Journal of Chemical Physics **2009**, *131*, 054106.

(267)   Guther, K.; Anderson, R. J.; Blunt, N. S.; Bogdanov, N. A.; Cleland, D.; Dattani, N.; Dobrautz, W.; Ghanem, K.; Jeszenszki, P.; Liebermann, N.; Manni, G. L.; Lozovoi, A. Y.; Luo, H.; Ma, D.; Merz, F.; Overy, C.; Rampp, M.; Samanta, P. K.; Schwarz, L. R.; Shepherd, J. J.; Smart, S. D.; Vitale, E.; Weser, O.; Booth, G. H.; Alavi, A. "NECI: N -Electron Configuration Interaction with an Emphasis on State-of-the-Art Stochastic Methods". Journal of Chemical Physics **2020**, *153*, 34107.

(268)   Spencer, J. S.; Blunt, N. S.; Choi, S.; Etrych, J.; Filip, M.-A.; Foulkes, W. M. C.; Franklin, R. S. T.; Handley, W. J.; Malone, F. D.; Neufeld, V. A.; Di Remigio, R.; Rogers, T. W.; Scott, C. J. C.; Shepherd, J. J.; Vigor, W. A.; Weston, J.; Xu, R.; Thom, A. J. W. "The HANDE-QMC Project: Open-Source Stochastic Quantum Chemistry from the Ground State Up". Journal of Chemical Theory and Computation **2019**, *15*, 1728–1742.

(269)   Brand, J. *Rimu.Jl*, RimuQMC, 2024.

(270)   Anderson, R. *M7*, 2024.

(271)   Dobrautz, W.; Weser, O.; Bogdanov, N. A.; Alavi, A.; Li Manni, G. "Spin-Pure Stochastic-CASSCF via GUGA-FCIQMC Applied to Iron–Sulfur Clusters". Journal of Chemical Theory and Computation **2021**, *17*, 5684–5703.

(272)   Weser, O.; Guther, K.; Ghanem, K.; Li Manni, G. "Stochastic Generalized Active Space Self-Consistent Field: Theory and Application". Journal of Chemical Theory and Computation **2022**, *18*, 251–272.

(273)   Dobrautz, W.; Smart, S. D.; Alavi, A. "Efficient Formulation of Full Configuration Interaction Quantum Monte Carlo in a Spin Eigenbasis via the Graphical Unitary Group Approach". The Journal of Chemical Physics **2019**, *151*, 094104.

(274)   Neufeld, V. A.; Thom, A. J. "Exciting Determinants in Quantum Monte Carlo: Loading the Dice with Fast, Low-Memory Weights". Journal of Chemical Theory and Computation **2019**, *15*, 127–140.

(275)   Li, J.; Otten, M.; Holmes, A. A.; Sharma, S.; Umrigar, C. J. "Fast Semistochastic Heat-Bath Configuration Interaction". The Journal of Chemical Physics **2018**, *149*, 214110.

(276) Weser, O.; Alavi, A.; Manni, G. L. "Exploiting Locality in Full Configuration Interaction Quantum Monte Carlo for Fast Excitation Generation". Journal of Chemical Theory and Computation **2023**, *19*, 9118–9135.

(277) Booth, G. H.; Smart, S. D.; Alavi, A. "Linear-Scaling and Parallelisable Algorithms for Stochastic Quantum Chemistry". Molecular Physics **2014**, *112*, 1855–1869.

(278) Blunt, N. S.; Rogers, T. W.; Spencer, J. S.; Foulkes, W. M. C. "Density-Matrix Quantum Monte Carlo Method". Physical Review B **2014**, *89*, 245124.

(279) Blunt, N. S.; Smart, S. D.; Kersten, J. A. F.; Spencer, J. S.; Booth, G. H.; Alavi, A. "Semi-Stochastic Full Configuration Interaction Quantum Monte Carlo: Developments and Application". THE JOURNAL OF CHEMICAL PHYSICS **2015**, *142*, 184107.

(280) Petruzielo, F. R.; Holmes, A. A.; Changlani, H. J.; Nightingale, M. P.; Umrigar, C. J. "Semistochastic Projector Monte Carlo Method". DOI: `10.1103/PhysRevLett.109.230201`.

(281) Spencer, J. S.; Blunt, N. S.; Foulkes, W. M. "The Sign Problem and Population Dynamics in the Full Configuration Interaction Quantum Monte Carlo Method". The Journal of Chemical Physics **2012**, *136*, 054110.

(282) Cleland, D.; Booth, G. H.; Alavi, A. "Communications: Survival of the Fittest: Accelerating Convergence in Full Configuration-Interaction Quantum Monte Carlo". Journal of Chemical Physics **2010**, *132*, 041103.

(283) Overy, C.; Booth, G. H.; Blunt, N. S.; Shepherd, J. J.; Cleland, D.; Alavi, A. "Unbiased Reduced Density Matrices and Electronic Properties from Full Configuration Interaction Quantum Monte Carlo". THE JOURNAL OF CHEMICAL PHYSICS **2014**, *141*, 244117.

(284) Haupt, J. P.; Hosseini, S. M.; López Ríos, P.; Dobrautz, W.; Cohen, A.; Alavi, A. "Optimizing Jastrow Factors for the Transcorrelated Method". The Journal of Chemical Physics **2023**, *158*, 224105.

(285) Feller, D.; Peterson, K. A.; Dixon, D. A. "A Survey of Factors Contributing to Accurate Theoretical Predictions of Atomization Energies and Molecular Structures". The Journal of Chemical Physics **2008**, *129*, 204105.

(286) Bytautas, L.; Ruedenberg, K. "Correlation Energy Extrapolation by Intrinsic Scaling. IV. Accurate Binding Energies of the Homonuclear Diatomic Molecules Carbon, Nitrogen, Oxygen, and Fluorine". The Journal of Chemical Physics **2005**, *122*, 154110.

(287) Harding, M. E.; Vázquez, J.; Ruscic, B.; Wilson, A. K.; Gauss, J.; Stanton, J. F. "High-Accuracy Extrapolated Ab Initio Thermochemistry. III. Additional Improvements and Overview". The Journal of Chemical Physics **2008**, *128*, 114111.

(288)   Dunning Jr., T. H. "Gaussian Basis Sets for Use in Correlated Molecular Calculations. I. The Atoms Boron through Neon and Hydrogen". The Journal of Chemical Physics **1989**, *90*, 1007–1023.

(289)   Sun, Q.; Berkelbach, T. C.; Blunt, N. S.; Booth, G. H.; Guo, S.; Li, Z.; Liu, J.; McClain, J. D.; Sayfutyarova, E. R.; Sharma, S.; Wouters, S.; Chan, G. K. L. "PySCF: The Python-based Simulations of Chemistry Framework". Wiley Interdisciplinary Reviews: Computational Molecular Science **2018**, *8*, e1340.

(290)   Hosseini, S. M. *Combining Orbital and Real Space Quantum Monte Carlo Methods*, doctoralThesis, 2024.

(291)   Ma, A.; Towler, M. D.; Drummond, N. D.; Needs, R. J. "Scheme for Adding Electron-Nucleus Cusps to Gaussian Orbitals". Journal of Chemical Physics **2005**, *122*, DOI: `10.1063/1.1940588`.

(292)   Umrigar, C. J.; Wilson, K. G.; Wilkins, J. W. "Optimized Trial Wave Functions for Quantum Monte Carlo Calculations". Physical Review Letters **1988**, *60*, 1719–1722.

(293)   Handy, N. "On the Minimization of the Variance of the Transcorrelated Hamiltonian". Molecular Physics **1971**, *21*, 817–828.

(294)   Nightingale, M. P.; Melik-Alaverdian, V. "Optimization of Ground- and Excited-State Wave Functions and van Der Waals Clusters". Physical Review Letters **2001**, *87*, 043401.

(295)   Toulouse, J.; Umrigar, C. J. "Optimization of Quantum Monte Carlo Wave Functions by Energy Minimization". The Journal of Chemical Physics **2007**, *126*, 084102.

(296)   Umrigar, C. J.; Toulouse, J.; Filippi, C.; Sorella, S.; Hennig, R. G. "Alleviation of the Fermion-Sign Problem by Optimization of Many-Body Wave Functions". Physical Review Letters **2007**, *98*, 110201.

(297)   Spink, G. G.; López Ríos, P.; Drummond, N. D.; Needs, R. J. "Trion Formation in a Two-Dimensional Hole-Doped Electron Gas". Physical Review B **2016**, *94*, 041410.

(298)   Ceperley, D. M. "The Statistical Error of Green's Function Monte Carlo". Journal of Statistical Physics **1986**, *43*, 815–826.

(299)   Becke, A. D. "A Multicenter Numerical Integration Scheme for Polyatomic Molecules". The Journal of Chemical Physics **1988**, *88*, 2547–2553.

(300)   Treutler, O.; Ahlrichs, R. "Efficient Molecular Numerical Integration Schemes". The Journal of Chemical Physics **1995**, *102*, 346–354.

(301)   Dobrautz, W.; Luo, H.; Alavi, A. "Compact Numerical Solutions to the Two-Dimensional Repulsive Hubbard Model Obtained via Nonunitary Similarity Transformations". Physical Review B **2019**, *99*, 075119.

(302)   Pfau, D.; Spencer, J. S.; Matthews, A. G. D. G.; Foulkes, W. M. C. "*Ab Initio* Solution of the Many-Electron Schrödinger Equation with Deep Neural Networks". Physical Review Research **2020**, *2*, 033429.

(303)   Hermann, J.; Schätzle, Z.; Noé, F. "Deep-Neural-Network Solution of the Electronic Schrödinger Equation". Nature Chemistry **2020**, *12*, 891–897.

(304)   Kutzelnigg, W.; Mukherjee, D. "Normal Order and Extended Wick Theorem for a Multiconfiguration Reference Wave Function". The Journal of Chemical Physics **1997**, *107*, 432–449.

(305)   Kutzelnigg, W.; Mukherjee, D. "Cumulant Expansion of the Reduced Density Matrices". The Journal of Chemical Physics **1999**, *110*, 2800–2809.

(306)   Haupt, J. P.; López Ríos, P.; Christlmaier, E. M. C.; Bogdanov, N.; Kats, D.; Alavi, A. "The Transcorrelated Method for Strongly Multireference Problems".

(307)   Le Roy, R. J.; Huang, Y.; Jary, C. "An Accurate Analytic Potential Function for Ground-State N2 from a Direct-Potential-Fit Analysis of Spectroscopic Data". The Journal of Chemical Physics **2006**, *125*, 164310.

(308)   Lyakh, D. I.; Musiał, M.; Lotrich, V. F.; Bartlett, R. J. "Multireference Nature of Chemistry: The Coupled-Cluster View". Chemical Reviews **2012**, *112*, 182–243.

(309)   Drummond, N. D.; Towler, M. D.; Needs, R. J. "Jastrow Correlation Factor for Atoms, Molecules, and Solids". Physical Review B **2004**, *70*, 235119.

(310)   Levine, B. G.; Durden, A. S.; Esch, M. P.; Liang, F.; Shu, Y. "CAS without SCF—Why to Use CASCI and Where to Get the Orbitals". The Journal of Chemical Physics **2021**, *154*, 090902.

(311)   Werner, H.-J.; and P. Celani, P. J. K.; Györffy, W.; Hesselmann, A.; Kats, D.; Knizia, G.; Köhn, A.; Korona, T.; Kreplin, D.; Lindh, R.; Ma, Q.; Manby, F. R.; Mitrushenkov, A.; Rauhut, G.; Schütz, M.; and T. B. Adler, K. R. S.; Amos, R. D.; Bennie, S. J.; Bernhardsson, A.; Berning, A.; Black, J. A.; Bygrave, P. J.; Cimiraglia, R.; Cooper, D. L.; Coughtrie, D.; Deegan, M. J. O.; Dobbyn, A. J.; Doll, K.; Dornbach, M.; Eckert, F.; Erfort, S.; Goll, E.; Hampel, C.; Hetzer, G.; Hill, J. G.; Hodges, M.; Hrenar, T.; Jansen, G.; Köppl, C.; Kollmar, C.; Lee, S. J. R.; Liu, Y.; Lloyd, A. W.; Mata, R. A.; May, A. J.; Mussard, B.; McNicholas, S. J.; Meyer, W.; Miller III, T. F.; Mura, M. E.; Nicklass, A.; O'Neill, D. P.; Palmieri, P.; Peng, D.; Peterson, K. A.; Pflüger, K.; Pitzer, R.; Polyak, I.; Reiher, M.; Richardson, J. O.; Robinson, J. B.; Schröder, B.; Schwilk, M.; Shiozaki, T.; Sibaev, M.; Stoll, H.; Stone, A. J.; Tarroni, R.; Thorsteinsson, T.; Toulouse, J.;

Wang, M.; Welborn, M.; Ziegler, B. *MOLPRO, Version , a Package of Ab Initio Programs*, Stuttgart, Germany.

(312) Werner, H.-J.; Knowles, P. J.; Knizia, G.; Manby, F. R.; Schütz, M. "Molpro: A General-Purpose Quantum Chemistry Program Package". Wiley Interdisciplinary Reviews: Computational Molecular Science **2012**, *2*, 242–253.

(313) Werner, H.-J.; Knowles, P. J.; Manby, F. R.; Black, J. A.; Doll, K.; Heßelmann, A.; Kats, D.; Köhn, A.; Korona, T.; Kreplin, D. A.; Ma, Q.; Miller III, T. F.; Mitrushchenkov, A.; Peterson, K. A.; Polyak, I.; Rauhut, G.; Sibaev, M. "The Molpro Quantum Chemistry Package". The Journal of Chemical Physics **2020**, *152*, 144107.

(314) Weser, O.; Liebermann, N.; Kats, D.; Alavi, A.; Li Manni, G. "Spin Purification in Full-CI Quantum Monte Carlo via a First-Order Penalty Approach". The Journal of Physical Chemistry A **2022**, *126*, 2050–2060.

(315) Loos, P.-F.; Scemama, A.; Blondel, A.; Garniron, Y.; Caffarel, M.; Jacquemin, D. "A Mountaineering Strategy to Excited States: Highly Accurate Reference Energies and Benchmarks". Journal of Chemical Theory and Computation **2018**, *14*, 4360–4379.

(316) Blunt, N. S.; Smart, S. D.; Booth, G. H.; Alavi, A. "An Excited-State Approach within Full Configuration Interaction Quantum Monte Carlo". THE JOURNAL OF CHEMICAL PHYSICS **2015**, *143*, 134117.

(317) Oddershede, J.; Grüner, N. E.; Diercksen, G. H. F. "Comparison between Equation of Motion and Polarization Propagator Calculations". Chemical Physics **1985**, *97*, 303–310.

(318) Huber, K. P.; Herzberg, G., *Constants of Diatomic Molecules* In *Molecular Spectra and Molecular Structure: IV. Constants of Diatomic Molecules*, Huber, K. P., Herzberg, G., Eds.; Springer US: Boston, MA, 1979, pp 8–689.

(319) Nielsen, E. S.; Jo/rgensen, P.; Oddershede, J. "Transition Moments and Dynamic Polarizabilities in a Second Order Polarization Propagator Approach". The Journal of Chemical Physics **1980**, *73*, 6238–6246.

(320) Ben-Shlomo, S. B.; Kaldor, U. "N2 Excitations below 15 eV by the Multireference Coupled-cluster Method". The Journal of Chemical Physics **1990**, *92*, 3680–3682.

(321) Simula, K.; López Ríos, P.; Christlmaier, E. M. C.; Haupt, J. P.; Alavi, A. "Effective Core Potentials in the Transcorrelated Method". **2025**, Submitted.

(322) Fournais, S.; Sørensen, T. Ø.; Hoffmann-Ostenhof, M.; Hoffmann-Ostenhof, T. "Non-Isotropic Cusp Conditions and Regularity of the Electron Density of Molecules at the Nuclei". Annales Henri Poincaré **2007**, *8*, 731–748.

(323)   Fournais, S.; Hoffmann-Ostenhof, M.; Hoffmann-Ostenhof, T.; Sørensen, T. Ø. "Sharp Regularity Results for Coulombic Many-Electron Wave Functions". Communications in Mathematical Physics **2005**, *255*, 183–227.

(324)   Tew, D. P. "Second Order Coalescence Conditions of Molecular Wave Functions". The Journal of Chemical Physics **2008**, *129*, 014104.

(325)   Filip, M.-A.; López Ríos, P.; Haupt, J. P.; Christlmaier, E. M. C.; Kats, D.; Alavi, A. "Transcorrelated methods applied to second row elements". The Journal of Chemical Physics **2025**, *162*, 064110.

(326)   Haupt, J. P.; López Ríos, P.; Christlmaier, E. M. C.; Filip, M.-A.; Simula, K.; Hauskrecht, J.; Liao, K.; Dobrautz, W.; Guther, K.; Cohen, A. J.; Alavi, A. *Pytchint: Transcorrelated Hamiltonian Integrals Library.*

(327)   Knowles, P. J.; Handy, N. C. "A Determinant Based Full Configuration Interaction Program". Computer Physics Communications **1989**, *54*, 75–83.

(328)   Fink, W. H. "Approach to Partially Predetermining Molecular Electronic Structure. The Li He Interaction Potential". The Journal of Chemical Physics **1972**, *57*, 1822–1826.

(329)   Baerends, E. J.; Ellis, D. E.; Ros, P. "Self-Consistent Molecular Hartree—Fock—Slater Calculations I. The Computational Procedure". Chemical Physics **1973**, *2*, 41–51.

(330)   Sachs, E. S.; Hinze, J.; Sabelli, N. H. "Frozen Core Approximation, a Pseudopotential Method Tested on Six States of NaH". The Journal of Chemical Physics **1975**, *62*, 3393–3398.

(331)   Message Passing Interface Forum *MPI: A Message-Passing Interface Standard Version 4.1*, Manual, 2023.

(332)   Dalcín, L.; Paz, R.; Storti, M. "MPI for Python". Journal of Parallel and Distributed Computing **2005**, *65*, 1108–1115.

(333)   Dalcín, L.; Paz, R.; Storti, M.; D'Elía, J. "MPI for Python: Performance Improvements and MPI-2 Extensions". Journal of Parallel and Distributed Computing **2008**, *68*, 655–662.

(334)   Dalcin, L. D.; Paz, R. R.; Kler, P. A.; Cosimo, A. "Parallel Distributed Computing Using Python". Advances in Water Resources **2011**, *34*, 1124–1139.

(335)   Dagum, L.; Menon, R. "OpenMP: An Industry Standard API for Shared-Memory Programming". IEEE Computational Science and Engineering **1998**, *5*, 46–55.

(336)   Blackford, L. S.; Petitet, A.; Pozo, R.; Remington, K.; Whaley, R. C.; Demmel, J.; Dongarra, J.; Duff, I.; Hammarling, S.; Henry, G., et al. "An Updated Set of Basic Linear Algebra Subprograms (BLAS)". ACM Transactions on Mathematical Software **2002**, *28*, 135–151.

(337)   Behnel, S. *The Cython Compiler for C-Extensions in Python*, Presented at EuroPython 2007, Vilnius, Lithuania, 2008.

(338)   Behnel, S.; Bradshaw, R.; Citro, C.; Dalcin, L.; Seljebotn, D. S.; Smith, K. "Cython: The Best of Both Worlds". Computing in Science & Engineering **2011**, *13*, 31–39.

(339)   Smith, K. W., *Cython: A Guide for Python Programmers*; "O'Reilly Media, Inc.": 2015.

(340)   Virtanen, P.; Gommers, R.; Oliphant, T. E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; van der Walt, S. J.; Brett, M.; Wilson, J.; Millman, K. J.; Mayorov, N.; Nelson, A. R. J.; Jones, E.; Kern, R.; Larson, E.; Carey, C. J.; Polat, İ.; Feng, Y.; Moore, E. W.; VanderPlas, J.; Laxalde, D.; Perktold, J.; Cimrman, R.; Henriksen, I.; Quintero, E. A.; Harris, C. R.; Archibald, A. M.; Ribeiro, A. H.; Pedregosa, F.; van Mulbregt, P.; SciPy 1.0 Contributors "SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python". Nature Methods **2020**, *17*, 261–272.

(341)   Wengert, R. E. "A Simple Automatic Derivative Evaluation Program". Commun. ACM **1964**, *7*, 463–464.

(342)   Bradbury, J.; Frostig, R.; Hawkins, P.; Johnson, M. J.; Leary, C.; Maclaurin, D.; Necula, G.; Paszke, A.; VanderPlas, J.; Wanderman-Milne, S.; Zhang, Q. *JAX: Composable Transformations of Python+NumPy Programs*, 2018.

The PyTCHInt Library

This appendix is based on the following software, to be released:

Haupt, J. P.; López Ríos, P.; Christlmaier, E. M. C.; Filip, M.-A.; Simula, K.; Hauskrecht, J.; Liao, K.; Dobrautz, W.; Guther, K.; Cohen, A. J.; Alavi, A. *Pytchint: Transcorrelated Hamiltonian Integrals Library*

The basis for how the library works is partially discussed in the following paper and its supplementary material:

Cohen, A. J.; Luo, H.; Guther, K.; Dobrautz, W.; Tew, D. P.; Alavi, A. "Similarity Transformation of the Electronic Schrödinger Equation via Jastrow Factorization". Journal of Chemical Physics **2019**, *151*, 061101

A.1. Introduction

All transcorrelated matrix element calculations presented in this dissertation have been performed using the group's `tchint` library or its Python extension, `pytchint`. This library works by either producing transformed integral files, dubbed `FCIDUMP` files for four-index integrals,[327] or `TCDUMP` files for the TC six-index integrals. Note that since the TC transformation is non-Hermitian, we do not have the same symmetries in these files as we do with conventional methods, or by interfacing with another program (such as `NECI`) or the Python interpreter (in the case of `pytchint`).

As described in section 2.5,

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

(A.1)

where

$$h_q^p = \langle p| h |q \rangle ,$$
$$V_{rs}^{pq} = \langle pq| r_{12}^{-1} |rs \rangle ,$$
$$K_{rs}^{pq} = \langle pq| \hat{K} |rs \rangle ,$$
$$L_{stu}^{pqr} = \langle pqr| \hat{L} |stu \rangle .$$

(A.2)

h_q^p and V_{rs}^{pq} are the one- and two-body terms from the electronic Schrödinger equation, familiar from conventional methods. Therefore, the key quantities to be evaluated by `tchint` are the non-Hermitian two-body integrals K_{rs}^{pq} and the Hermitian three-body integrals L_{stu}^{pqr}.

Once these integrals are evaluated, as long as we take care about using the correct bra and ket (a detail not important for Hermitian problems), many different methodologies can be applied with these integrals, such as the frozen-core approximation,[328–330] or xTC.[190]

A.2. Matrix Element Evaluation

The matrix elements not present in conventional methods are K_{rs}^{pq} and L_{stu}^{pqr}. For these, we integrate on a grid, typically Treutler-Ahlrichs integration grids,[299,300] which are atom-centred grids commonly used in density functional theory. These grids are obtained via `pyscf`.

The two-electron matrix elements we need to evaluate are:

$$K_{rs}^{pq(1)} = \langle pq | \nabla_1 u(\boldsymbol{r}_1, \boldsymbol{r}_2) \cdot \nabla_1 | rs \rangle \tag{A.3}$$

$$K_{rs}^{pq(2)} = \langle pq | \nabla_1^2 u(\boldsymbol{r}_1, \boldsymbol{r}_2) | rs \rangle \tag{A.4}$$

$$K_{rs}^{pq(3)} = \langle pq | (\nabla_1 u(\boldsymbol{r}_1, \boldsymbol{r}_2))^2 | rs \rangle . \tag{A.5}$$

Discretising on a grid of N_{grid} points with weights w, we have

$$K_{rs}^{pq(1)} = \sum_{mn}^{N_{\text{grid}}} \phi_p(\boldsymbol{r}_m) \nabla_{\boldsymbol{r}_m} \phi_r(\boldsymbol{r}_m) \cdot \nabla_{\boldsymbol{r}_m} u(\boldsymbol{r}_m, \boldsymbol{r}_n) \phi_q(\boldsymbol{r}_n) \phi_s(\boldsymbol{r}_n) w(\boldsymbol{r}_m) w(\boldsymbol{r}_n). \tag{A.6}$$

Naive integration of this value yields $\mathcal{O}(N_{\text{grid}}^2 M^4)$ performance (where M is the number of basis functions). However, we may improve this by first integrating over one coordinate and storing the intermediate value. Moreover, for $K_{rs}^{pq(2)}$, it is more efficient to integrate by parts. Therefore, for the full K matrix, we calculate the intermediate value

$$X_s^q(\boldsymbol{r}_2) = \int \mathrm{d}^3 r_1 \ \nabla_1 u(\boldsymbol{r}_1, \boldsymbol{r}_2) \cdot [\phi_p(\boldsymbol{r}_1) \nabla_1 \phi_r(\boldsymbol{r}_1) - \phi_r(\boldsymbol{r}_1) \nabla_1 \phi_p(\boldsymbol{r}_1)] \tag{A.7}$$

$$+ \int \mathrm{d}^3 r_1 \ \phi_p(\boldsymbol{r}_1) (\nabla u(\boldsymbol{r}_1, \boldsymbol{r}_2))^2 \phi_r(\boldsymbol{r}_1). \tag{A.8}$$

We can then obtain the K matrix by

$$K_{rs}^{pq} = \int \mathrm{d}^3 r_2 \ \phi_q(\boldsymbol{r}_2) X_s^q(\boldsymbol{r}_2) \phi_s(\boldsymbol{r}_2) \tag{A.9}$$

for a cost of $\mathcal{O}(N_{\text{grid}}^2 M^2 + N_{\text{grid}} M^4)$.

The three-body matrix,

$$L_{stu}^{pqr} = \langle pqr| \nabla_1 u(\boldsymbol{r}_1, \boldsymbol{r}_2) \cdot \nabla u_1(\boldsymbol{r}_1, \boldsymbol{r}_3) |stu\rangle \tag{A.10}$$

may similarly be resolved by calculating the intermediate value

$$\boldsymbol{Y}_{qt}(\boldsymbol{r}_1) = \int \mathrm{d}^3 r_2 \ \phi_q(\boldsymbol{r}_2) \nabla_1 u(\boldsymbol{r}_1, \boldsymbol{r}_2) \phi_t(\boldsymbol{r}_2) \tag{A.11}$$

to give

$$L_{stu}^{pqr} = \int \mathrm{d}^3 r_1 \ \phi_p(\boldsymbol{r}_1) \boldsymbol{Y}_{qt}(\boldsymbol{r}_1) \cdot \boldsymbol{Y}_{ru}(\boldsymbol{r}_1) \phi_s(\boldsymbol{r}_1). \tag{A.12}$$

These integrations are simply parallelisable, and `tchint` takes advantage of this by leveraging distributed memory with the Message Passing Interface (MPI)[331–334] standard and the OpenMP[335] application programming interface (API). We also make liberal use of BLAS[336] routines for further performance gains. Moreover, we take advantage of the fact that L is Hermitian, substantially reducing storage requirements. We may also store L sparsely since many elements are small, or use the xTC approximation to bypass storing L altogether and instead calculating modified four-index integrals and outputting a non-Hermitian `FCIDUMP`.

A.3. Interface

`tchint` is predominantly written in Fortran, but most of the important interfacing subroutines (such as returning given matrix elements) have C bindings. This allows `tchint` to be easily interfaced with other libraries, notably for TC-FCIQMC in `NECI`[267] and `M7`.[270] By using this interface, these programs may be used to perform transcorrelated calculations, with all the features present in `tchint`.

`tchint` has also been given a Python interface called `pytchint` by making use of the Cython language extension.[337–339] Cython features C-like performance with easy integration in Python to allow for interactive computing and interfacing with the vast scientific package ecosystem available in Python. Python is typically also easier for rapid prototyping, a crucial feature in the world of scientific research, while Cython provides efficiency and a comprehensive profiling toolset.

With `pytchint`, a user might interactively work with transcorrelated integrals by first calculating them,

```
import pytchint
options = pytchint.options.TcOptions(yaml_input="tchint.yml")
options.eval_mode = "xtc"
solver = pytchint.solver.Solver(options=options)
solver.run()
```

and then e.g. query individual matrix elements calculated via Slater-Condon rules with `solver.sltcnd` or return the variance of the reference with `solver.refconn`.

A.3.1. Deterministic Optimisation

One of the major motivations for developing the Python interface is removing VMC from the Jastrow optimisation pipeline, as this can be a particularly expensive step, especially for larger systems. Thanks to the Python interface, we can interact directly with Python's large ecosystem of optimisation packages, such as SciPy.[340] This allows for rapid prototyping and ease of development.

For each optimisation step, we pass the Jastrow parameters in Python, then calculate the relevant integrals for the variance of the reference (see chapter 4) by passing them to Fortran, and update the parameters according to an optimisation algorithm by passing the parameters back to Python. All that is needed to be added are gradients for the variance of the reference with respect to Jastrow factor parameters. As the Jastrow factors are linear in their parameters, this is a tractible problem.

However, as calculations are done entirely in `pytchint`, we are restricted to a basis set and do not enjoy the CBS optimisation as we do in continuum VMC. As an example, we consider calculations with the Li atom. Consider, moreover, the simple parameter-free electron-electron-cusp-correcting Jastrow factor

$$J_0 = \frac{1}{2}\bar{r}_{ij} \tag{A.13}$$

with the additional terms for the electron-nucleus cusp described in chapter 4. Here, $\bar{r}_{ij} = 1 - \exp(-\alpha r_{ij})/\alpha$ with $\alpha = 2.0$. Consider two possible additional terms, each with one variational parameter,

$$J_{ee} = c_{ee}\bar{r}_{ij}^2 \tag{A.14}$$

$$J_{en} = c_{en}\bar{r}_i^2 \tag{A.15}$$

where for the electron-nucleus term $\bar{r}_i = 1 - \exp(-\beta r_{ij})/\beta$ with $\beta = 2.0$.

Reference energies and variance of the reference values are presented in figure A.1. As illustrated, the minimum may shift based on basis set, so a careful selection is necessary for obtaining the optimal Jastrow factor.

Preliminary results are promising, and this may prove an interesting alternative to VMC, or a way of "correcting" smaller VMC calculations, to improve efficiency as well as reproducibility.
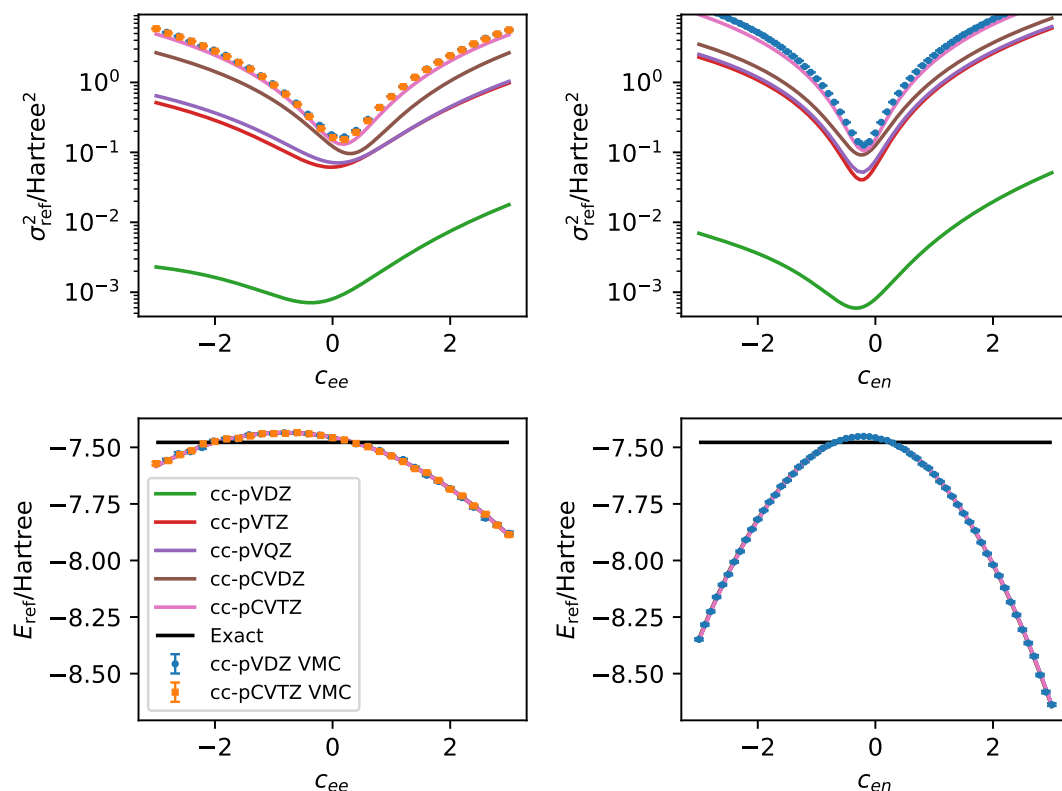
Figure A.1. Variance of the reference and reference energy values (top and bottom panels, respectively) with a single-parameter Jastrow factor, for the Li atom with the cc-pVXZ basis sets, with $X =$D, T, Q, as well as basis sets core-valence correlation at the double- and triple-zeta level. The left panel introduces an electron-electron term J_{ee} whereas the right panel introduces an electron-nucleus term J_{en}. The location of the minimum may be different compared to that obtained by continuum VMC, resulting in a suboptimal Jastrow factor. Data courtesy of Maria-Andreea Filip, who is leading this investigation. Results are preliminary and are intended to be presented in a future publication.

A.4. Parallel Efficiency

As a benchmark, we calculate N_2 with the aug-cc-pVQZ basis (14 electrons in 160 orbitals) with the xTC approximation. Elapsed time is plotted as a function of the number of compute cores used in figure A.2, with favourable scaling. This illustrations the parallel performance of the `pytchint` library.

Moreover, thanks to the use of Cython, performance when directly running Fortran when compared to running via Python is roughly identical (468 seconds or 465 seconds with 128 cores for Fortran and Python, respectively).
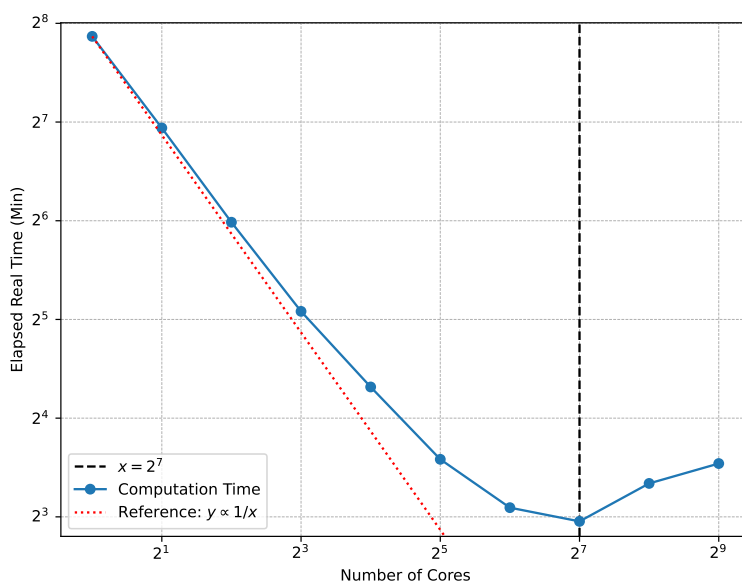

Figure A.2. Elapsed real time (or walltime) of calculating all-electron transcorrelated integrals under the xTC approximation for N_2 with the aug-cc-pVQZ basis. Note the logarithmic axes. Performance improves drastically when given additional cores, highlighting the parallel capabilities of the `pytchint` library. The dotted red line indicates ideal scaling. The dotted vertical line at 2^7 cores indicates additional nodes. Introducing more nodes, walltime slightly increases, indicating the additional MPI overhead of internode communication is not sufficiently compensated for by the additional cores for this size problem. However, it is worth noting that due to the distributed memory model, these calculations offer similar time performance but the memory load per node is reduced by roughly a factor equal to the number of nodes. Calculations were performed on AMD EPYC 9554 64-Core processors, and each node used 128 cores.

A.5. Conclusion and Outlook

Our in-house transcorrelated integral evaluation library is already flexible and performant, and is already used for most TC-related publications from the group. For wider adoption,

we plan to improve documentation and ease of use through the Python interface, and have a form of Jastrow optimisation or (quasi-)universal Jastrow form such as those described in chapter 6. This way, `pytchint` is a standalone library and would not rely on external software.

Additional features we intend to investigate in the near future are: parallelism with graphics processing units (as our integrals are easily parallelisable), which is already being investigated in the context of TC, further support for periodic solids, and implementing more of the code in high-performance Python to allow easier use of additional features such as automatic differentiation[341] with JAX,[342] which would allow support for more complex Jastrow factors without the need for manually calculating analytic gradients.

Academic Curriculum vitae

Education

2020 − 2025   PhD student, Theoretical Chemistry
Max Planck Institute for Solid State Research, Stuttgart, Germany
Thesis: Development of the Transcorrelated Full Configuration
  Interaction Quantum Monte Carlo Method
Advisor: Prof Ali Alavi

2014 − 2019   Bachelor of Science, Honours Physics and Mathematics, Minor in
  Computer Science with Co-operative Education
University of British Columbia, Vancouver, Canada
Thesis: Effects of Anisotropic Long-Range Hopping on Anderson
  Localization
Advisor: Prof Roman Krems

Publications

First Author

(1) Haupt, J. P.; Hosseini, S. M.; López Ríos, P.; Dobrautz, W.; Cohen, A.; Alavi, A. "Optimizing Jastrow Factors for the Transcorrelated Method". The Journal of Chemical Physics **2023**, *158*, 224105

In Preparation

(2) Haupt, J. P. et al. "The Transcorrelated Method for Strongly Multireference Problems"
(3) Haupt, J. P. et al. *Pytchint: Transcorrelated Hamiltonian Integrals Library*
(4) Haupt, J. P. et al. "Modular Construction of Jastrow Factors for the Transcorrelated Method"

Coauthor

(5) Filip, M.-A.; López Ríos, P.; Haupt, J. P.; Christlmaier, E. M. C.; Kats, D.; Alavi, A. "Transcorrelated methods applied to second row elements". The Journal of Chemical Physics **2025**, *162*, 064110
(6) Simula, K.; López Ríos, P.; Christlmaier, E. M. C.; Haupt, J. P.; Alavi, A. "Effective Core Potentials in the Transcorrelated Method". **2025**, Submitted

In Preparation

(7) Christlmaier, E. M. C. et al. "Lifting the spin-restriction in transcorrelation"
(8) Filip, M.-A. et al. "Deterministic Jastrow Factor Optimization for the Transcorrelated Method"