MONTE CARLO EXPLICITLY CORRELATED METHODS

BY

COLE JOHNSON

DISSERTATION

Submitted in partial fulfillment of the requirements for the degree of Doctor of Philosophy in Chemistry in the Graduate College of the University of Illinois at Urbana-Champaign, 2018

Urbana, Illinois

Doctoral Committee:

Professor So Hirata, Chair
Professor David Ceperley
Professor Nancy Makri
Assistant Professor Lucas Wagner
Abstract

Solving the non-relativistic time-independent electronic Schrödinger equation is in general difficult and requires approximation. For experimental accuracy, wave-function based methods require a large set of basis functions and inclusion of instantaneous correlation through expensive correlated methods. The methods that have been developed to account for the incompleteness of the basis set, the R12/F12 methods, create high dimensional integrals that need to be separated with the resolution of the identity, are limited in their form of the correlation factor due to analytical integration, and not highly parallel scalable. The solution to these drawbacks proposed in this work is Monte Carlo (MC).

The stochastic second-order many-body perturbation theory, or the MC-MP2-F12 method, was developed for highly parallel evaluation of second-order many-body perturbation theory (MP2) energies near the complete basis set (CBS) limit. Single molecule energies were on average closer to the CBS limit than the corresponding method with a much larger basis set. Many different reaction energies for small molecules were computed showing a mean error from the CBS limit result within chemical accuracy. Two different methods were used the full variational MP2-F12 correction, MC-MP2-F12(VBX), and a non-variational approximate form only satisfied at the minimum of the MC-MP2-F12(VBX) formula, MC-MP2-F12(V). Despite previous assumptions, the MC-MP2-F12(V) formula is accurate not only for absolute energies but relative energies as well. Scaling for relative errors was shown to be $O(n^4)$ where $n$ is the number of basis functions, one order lower than the corresponding deterministic method. Due to the MC-MP2-F12(V) and more complete MC-MP2-F12(VBX) having the same asymptotic scaling as $n$ increases, it is generally recommended that one use the VBX method for larger molecules. Various correlation factors were tested but the Slater-type geminal (STG) developed by Ten-no was confirmed to be the best.

A more extensive study of different functional forms of correlation factors was conducted using the MC-MP2-F12 method with a total of 17 correlation factors in order to elucidate qualities of the correlation hole and shape. Higher-order cusp conditions, or derivatives of the wavefunction, and their properties were also studied. It was found that every correlation factor that had the best convergence to the CBS limit had a very specific shape on the range of 0 to 1.5 Bohr. Despite having vastly differing long-range behavior, the best correlation factors gave very similar energies. This was found to be due to the decoupling of electrons at long distance, and the dominance of the orbital expansion at large inter-electron distance, $r_{12}$. While the importance of satisfying the cusp
condition at $r_{12} = 0$ could not be determined, the study confirmed that the intermediate region is of the most importance in general.

Lastly, the MC-F12 algorithm developed for MC-MP2-F12 was extended to explicitly correlated second-order Green’s function theory (GF2-F12) for basis-set corrected ionization potentials (IPs). The same set of benchmark organic molecules that were studied in the original GF2-F12 study were compared to verify the usefulness of the MC algorithm. Analogous to MC-MP2-F12, the two different methods MC-GF2-F12($V$) and MC-GF2-F12($V_{BX}$) were tested. A mean average error of 0.049 eV and 0.018 eV was achieved for the MC-GF2-F12($V$) and MC-GF2-F12($V_{BX}$) methods respectively. System size scaling was found to be $O(n^4)$. As a demonstration of size scalability, the first IPs of fullerenes $C_{60}$ and $C_{70}$ were corrected from HF at the MC-GF2-F12($V_{BX}$) level. Errors of 0.37 eV and 0.05 eV from experiment were achieved for $C_{60}$ and $C_{70}$ respectively. The sources of the large error in $C_{60}$ is unknown. Further accuracy can be expected from developing the full non-diagonal frequency-dependent formalism with MC, as well as combining MC-GF2-F12 with the MC-GF3 and MC-GF4 methods.
Acknowledgements

This would not have been possible without many people. Firstly, I would like to acknowledge my advisor and mentor, So Hirata. I have learned more than I ever could have imagined and produced some amazing work with his help and guidance. I am extremely grateful for the self-motivated environment that he cultivated in his lab.

It has also been a great experience sharing this lab with many enthusiastic and intelligent colleagues. Former group members, Misha, Matt, and Ryan demonstrated a daunting level of knowledge and technique that inspired me to improve and establish myself as a prominent researcher in the lab. I would like to especially thank Misha for all of his patience and constant assistance during my first year. I would not have been able to accomplish what I did in this time without him. Newer lab members, Alex and Evan have provided a fun and exciting environment to do research being open to collaborating on programming techniques and discussing general theories. I would also especially like to thank Alex for his abundant expertise in programming and helping me throughout the process with my many questions.

At my undergraduate university, University of St. Thomas, I would like to especially thank Dr. Gary Mabbott and Dr. Joseph Brom. If Gary would not have persuaded me to switch into computational research from experimental I probably never would have had the opportunity to accomplish the things I have up until now. And Joe, for being willing to take me under his wing and teach me all he had to know about computational chemistry. His passion for teaching was infectious and incredibly inspiring. I would also like to thank Dr. Magda Stolarska for her support in my love of mathematics and pushing me to pursue more advanced math which has immensely helped in understanding and developing new ideas in my research.

Most importantly, I must thank all of my friends and family who have shaped my life over the years. My mom has always been a positive force in my life and has always been my number one fan. I would like to thank my lovely girlfriend Cara. I would not be the person I am today without her. Her endless support, patience, and honesty with me has been invaluable during this whole process. I am incredibly lucky to have all of these great people in my life and I can’t wait to see where we all end up.
# Table of Contents

## List of Abbreviations

<table>
<thead>
<tr>
<th>Chapter 1</th>
<th>Introduction</th>
<th>1</th>
</tr>
</thead>
<tbody>
<tr>
<td>1.1</td>
<td>Background</td>
<td>1</td>
</tr>
<tr>
<td>1.2</td>
<td>MP2 Formalism</td>
<td>5</td>
</tr>
<tr>
<td>1.3</td>
<td>Original MP2-R12 formulation</td>
<td>6</td>
</tr>
<tr>
<td>1.3.1</td>
<td>Correlation cusp condition</td>
<td>6</td>
</tr>
<tr>
<td>1.3.2</td>
<td>Klopper and Kutzelnigg</td>
<td>6</td>
</tr>
<tr>
<td>1.4</td>
<td>Improvements to MP2-R12</td>
<td>8</td>
</tr>
<tr>
<td>1.4.1</td>
<td>Improved correlation factors</td>
<td>8</td>
</tr>
<tr>
<td>1.4.2</td>
<td>Fixed amplitudes</td>
<td>8</td>
</tr>
<tr>
<td>1.5</td>
<td>Metropolis Monte Carlo with importance sampling</td>
<td>9</td>
</tr>
<tr>
<td>1.6</td>
<td>Redundant walker algorithm</td>
<td>10</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 2</th>
<th>MC-MP2-F12-V/VBX method</th>
<th>12</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Introduction</td>
<td>12</td>
</tr>
<tr>
<td>2.2</td>
<td>Theory</td>
<td>13</td>
</tr>
<tr>
<td>2.2.1</td>
<td>$E_{\text{F12}}$</td>
<td>13</td>
</tr>
<tr>
<td>2.2.2</td>
<td>$E_{\text{V F12}}$</td>
<td>18</td>
</tr>
<tr>
<td>2.2.3</td>
<td>$E_{\text{BX F12}}$</td>
<td>20</td>
</tr>
<tr>
<td>2.3</td>
<td>Monte Carlo algorithm</td>
<td>27</td>
</tr>
<tr>
<td>2.3.1</td>
<td>$E_{\text{V F12}}$</td>
<td>28</td>
</tr>
<tr>
<td>2.3.2</td>
<td>$E_{\text{BX F12}}$</td>
<td>29</td>
</tr>
<tr>
<td>2.3.3</td>
<td>Statistical uncertainty</td>
<td>31</td>
</tr>
<tr>
<td>2.4</td>
<td>Results and discussion</td>
<td>32</td>
</tr>
<tr>
<td>2.4.1</td>
<td>Geminals and V versus VBX formulas</td>
<td>33</td>
</tr>
<tr>
<td>2.4.2</td>
<td>Correlation energies</td>
<td>37</td>
</tr>
<tr>
<td>2.4.3</td>
<td>Reaction energies</td>
<td>39</td>
</tr>
<tr>
<td>2.4.4</td>
<td>System-size scaling</td>
<td>44</td>
</tr>
<tr>
<td>2.4.5</td>
<td>Computer-size scaling</td>
<td>47</td>
</tr>
<tr>
<td>2.5</td>
<td>Conclusions</td>
<td>49</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Chapter 3</th>
<th>Explicit correlation factors</th>
<th>53</th>
</tr>
</thead>
<tbody>
<tr>
<td>3.1</td>
<td>Introduction</td>
<td>53</td>
</tr>
<tr>
<td>3.2</td>
<td>Computational details</td>
<td>54</td>
</tr>
<tr>
<td>3.3</td>
<td>Results</td>
<td>56</td>
</tr>
<tr>
<td>3.4</td>
<td>Discussion</td>
<td>60</td>
</tr>
<tr>
<td>3.4.1</td>
<td>Cusp ($r_{12} = 0$)</td>
<td>60</td>
</tr>
</tbody>
</table>
# List of Abbreviations

<table>
<thead>
<tr>
<th>Abbreviation</th>
<th>Meaning</th>
</tr>
</thead>
<tbody>
<tr>
<td>IP</td>
<td>Ionization potential</td>
</tr>
<tr>
<td>EA</td>
<td>Electron attachment</td>
</tr>
<tr>
<td>MP2</td>
<td>Many-body second-order perturbation theory.</td>
</tr>
<tr>
<td>CBS</td>
<td>Complete basis set.</td>
</tr>
<tr>
<td>CC</td>
<td>Coupled cluster.</td>
</tr>
<tr>
<td>HF</td>
<td>Hartree–Fock.</td>
</tr>
<tr>
<td>MBPT</td>
<td>Many-body perturbation theory.</td>
</tr>
<tr>
<td>MBGF</td>
<td>Many-body Green’s function theory.</td>
</tr>
<tr>
<td>AO</td>
<td>Atomic orbital</td>
</tr>
<tr>
<td>MO</td>
<td>Molecular orbital.</td>
</tr>
<tr>
<td>MC</td>
<td>Monte Carlo.</td>
</tr>
<tr>
<td>RI</td>
<td>Resolution of the identity.</td>
</tr>
<tr>
<td>DFT</td>
<td>Density functional theory.</td>
</tr>
<tr>
<td>GTO</td>
<td>Gaussian-type orbital.</td>
</tr>
<tr>
<td>ABS</td>
<td>Auxiliary basis set.</td>
</tr>
<tr>
<td>CPU</td>
<td>Central processing unit.</td>
</tr>
<tr>
<td>GPU</td>
<td>Graphical processing unit.</td>
</tr>
<tr>
<td>QMC</td>
<td>Quantum Monte Carlo.</td>
</tr>
<tr>
<td>R12/F12</td>
<td>Explicitly correlated methods.</td>
</tr>
<tr>
<td>BGQMC</td>
<td>Brueckner–Goldstone quantum Monte Carlo.</td>
</tr>
</tbody>
</table>
Chapter 1: Introduction

1.1 Background

One goal of quantum chemistry is to solve the non-relativistic time-independent Schrödinger equation for electrons

\[ \hat{H} \Psi = E \Psi, \]

where \( \hat{H} \) is the Hamiltonian operator, \( E \) is the energy of the system, and \( \Psi \) is the wavefunction that describes that system.\(^1\) This equation governs the interactions between electrons, the energies of their states, and the energy of the system as a whole. The Schrödinger equation cannot in general be solved exactly, and therefore it is necessary to develop approximate ways to solve the equation. The two main ways the Schrödinger equation is approximated are through so-called wavefunction based methods or density functional theory (DFT).\(^2,3,4,5\) This thesis will focus on the wave-function based methods. The approximations within the wave-function based methods involve the choice of basis set and the level of correlated treatment.

If the exact wavefunction is always known, the exact system and energy of that system could be solved for and none of these approximate schemes would be necessary. In general, this is not the case, and the wavefunction is approximated with what is known as a basis set. A basis set is a set of functions chosen to represent each atom, and the functions are used to construct the orbitals which describe the probability distribution of each electron. The size of the basis set limits the flexibility of the orbitals we can construct to minimize the energy. A set of infinite functions offers infinite flexibility and is known as the complete basis set (CBS) limit. Dunning et al.\(^6\) developed a set of basis functions that systematically converge to the CBS limit as the angular momentum of the functions, \( L \), is increased. As \( L \) increases, the number of functions increases rapidly. This is known as the basis set incompleteness problem: in order to correct for the finite size of our basis set, larger basis sets must be used. While increasing the basis set increases accuracy, this process suffers from extremely diminishing returns. As discussed later, perturbation approximations to solve the Schrödinger equation have high polynomial scaling on the number of basis functions, which is problematic when approaching the CBS limit.

Wave-function based methods start from the Hartree-Fock (HF) approximation and a chosen basis set, which is a mean-field solution to the Schrödinger equation in that basis. In HF, each
electron accounts for the mean-field interaction of other electrons instead of the instantaneous coulomb repulsion. However, HF accounts approximately but not fully for spin-interaction as it prevents two electrons of the same spin to be found at the same point in space. This is a cheap and effective starting approximation, recovering more than 99% of the total energy in many cases, but turns out to be lacking in accuracy when comparing to experimental results. Even with a very large basis set, HF in many cases does not achieve what is called “chemical accuracy”, or kT accuracy to experimental results. Chemical accuracy is a benchmark for theoretical methods to match or exceed experimental accuracy, which is around 1 kcal/mol. If chemical accuracy is desired then it is necessary to go beyond the HF approximation, and use a large basis set, both of which increase the cost of using the method.

Post-HF methods were developed to correct the HF approximation for it’s lack of electron correlation. The two post-HF methods studied here are many-body perturbation theory (MBPT)\(^7;8;9;10\) and many-body Green’s function theory (MBGF)\(^11;12;13\). These methods both include higher orders of electron interactions to be accounted for in the wavefunction and corresponding energy. The higher the order of the interactions, the more expensive the methods. Both methods utilized in this study, second-order many-body perturbation theory (MP2) and second-order many-body Green’s function theory (GF2), scale as \(O(n^5)\), where \(n\) is the size of the basis set, due to an atomic orbital (AO) to molecular orbital (MO) transformation. The two-electron integrals describing the two-electron interactions become difficult to calculate and store as the system size becomes large.

The focus of this study is not on improving the scaling of the post-HF methods, but improving the convergence of the basis set. If it were possible to use a small basis set and still get close to chemically accurate results, these post-HF methods could be used on larger systems. The basis functions functions developed by Dunning et al.\(^6\) are only functions of one electron coordinate. Slater and Hylleraas in the late 1920’s\(^14;15;16\) found that a significantly more accurate solution for the energy of He could be obtained if basis functions based on the inter-electron distance, denoted \(r_{12}\), were added to the regular expansion. Kato\(^17\) confirmed Slater and Hylleraas, and found that as two-electrons coalesce, the wavefunction becomes linear in \(r_{12}\), suggesting possibly that the one-electron basis was inadequate in this region. To describe electrons with inherently correlated behavior, it is reasonable that basis functions dependent on the inter-electron distance would dramatically increase the quality of our basis set. More specifically, one-electron basis functions are a poor choice when trying to simulate the cusp in the wavefunction where two electrons coalesce. The function of \(r_{12}\) used in our explicitly correlated basis is known as the correlation factor or geminal. The next step was to add these “explicitly correlated” or “geminal” basis functions into the post-HF methods to derive correlated energies and other important chemical properties at a much
cheaper cost.

The first development of MP2-R12, which is the MP2 method combined with the R12 basis functions, was developed by Klopper and Kutzlenigg. They showed that by using a linear correlation factor of \( r_{12} \), one could increase the basis set convergence of MP2 from \( O(L^{-3}) \) to \( O(L^{-7}) \). This dramatically reduces the size of the basis set required for chemically accurate results. Many people have since improved the MP2-R12 method by increasing the dependence of the basis functions on \( r_{12} \), the most notable and widely used being the slater-type geminal (STG) developed by Ten-no. This more complicated dependence on \( r_{12} \) more accurately mimics the decay of the electron density as two-electrons coalesce. These new methods with more advanced correlation factors are called the MP2-F12 method. It has since been extended to the more advanced coupled cluster (CC) methods and much more recently MBGF.

The F12 ansatz is, however, not without its drawbacks. Firstly, integrals of a geminal multiplying Gaussian-type orbitals (GTOs) are evaluated analytically, limiting the forms of such factors to a handful of the simplest ones. Second, in the derivation of MP2-R12 an orthogonality projector is required when introducing the new explicit functions of \( r_{12} \) to prevent over-counting from the regular pair excitations by the geminal basis. When the orthogonal projector is expanded, numerous three-electron (9-dimensional) and four-electron (12-dimensional) integrals involving both occupied and virtual orbitals emerge. The cost of their evaluation would be prohibitive. To alleviate the cost of evaluating these high-dimensional integrals, a resolution of the identity (RI) is used to separate them into a large number of two-electron integrals. In addition to creating many intermediate two-electron integrals that need to be evaluated, the RI must be converged requiring a large basis set also called the auxiliary basis set (ABS), defeating some of the original purpose of using F12 in the first place. And lastly, these F12 integrals together with the ordinary two-electron integrals for MP2, need to be pre-calculated, stored, and used in a series of dense matrix multiplications. Such computational steps are poorly scalable with system size, due to the high orbital scaling of MP2. They are also non-scalable with respect to the number of processors due to frequent inter-processor communications interrupting the underlying parallelism.

The solution this study proposes to all of these setbacks is to use Monte Carlo (MC), specifically Metropolis MC, to evaluate such integrals. MC is a well established way to sample high-dimensional integrals when quadrature is not a reasonable option. Using randomly generated numbers, Metropolis MC generates a sequence of points through MC "moves" or "steps" that when summed up systematically converge to the exact value of the integral. MC has already been used as a way to calculate quantum properties from the QMC community. Given that the MC algorithms only require the form of the integrand, and a distribution to sample
the integral, almost any correlation factor may be used. There are some contingencies required
that will be discussed later in Chapter 3 specific to the developed MC-MP2-F12 methodology.
Secondly, we utilize the advantage of MC when sampling a high dimensional space to evaluate
the high dimensional integrals directly avoiding the use of RI and the large basis set to converge
it. And lastly, MC falls into the category of highly parallel algorithms meaning multiple MC
runs are done in parallel and then averaged for a further converged result. An MC version of
the MP2 method, where the two-electron integrals of MP2 are evaluated directly by Monte-Carlo
has already been developed,\textsuperscript{43} The algorithms present in the MC-MP2 method were then later
extended to a simplified version of MC-MP2-F12 for only a part of the full MP2-F12 expression,
called MC-MP2-F12(V),\textsuperscript{44} This established that the integrals generated from the R12 operators
could in fact be integrated with a similar MC algorithm as used in MC-MP2. We continue these
studies of MC methods to develop the full explicitly correlated equation including all terms, termed
the MC-MP2-F12(VBX) method. In addition, we develop the MC analogue of the newly derived
GF2-F12.

In Chapter 2, the Monte Carlo explicitly correlated second-order perturbation theory or MC-
MP2-F12(VBX) formalism is developed and details of the method are explored. Reaction energies
are quantified and scaling on number of processors and system size are analyzed. In Chapter 3, the
same methodology developed in Chapter 2 is used to explore a large number of correlation factors
never attempted before in the F12 community. Insights about the cusp condition and what qualities
make the best performing correlation factor are given justified by calculations on H2O and CH4.
In Chapter 4, an analogue of the MC-MP2-F12(VBX) method for ionization energies based on
the so-called Green’s function methods termed the MC-GF2-F12 method is derived. Agreement
with deterministic results and some large scale results on fullerenes are demonstrated. Finally,
in Chapter 5 we will conclude the advantages and disadvantages of the new explicitly correlated
MC methods. Comparisons and validations will be made with the corresponding deterministic
methods and future implementations will be detailed. The rest of this chapter will discuss the
MP2 formalism, the details of how the R12 basis is included, the improvements to MP2-R12, how
Metropolis MC is done, and a short introduction about the “redundant walker” algorithm used
throughout this thesis.
1.2 MP2 Formalism

The MP2 energy is defined through the first order wavefunction, denoted $|\Psi^{(1)}\rangle$, as a sum of doubly excited determinants

$$|\Psi^{(1)}\rangle = \sum_{i<j}^{\text{occ.}} \sum_{a<b}^{\text{vir.}} t_{ij}^{ab} \Phi_{ij}^{ab},$$

(1.1)

where $i, j, ..$ are occupied orbitals, $a, b, ...$ are virtual orbitals, $t_{ij}^{ab}$ is the excitation amplitude for exciting the two-electrons from occupied pair $ij$ to virtual pair $ab$, $\Phi_{ij}^{ab}$ is the corresponding doubly excited Slater determinant, and $|\Psi^{(1)}\rangle$ is the MP2 wavefunction. The second-order corrected energy is calculated from the $|\Psi^{(1)}\rangle$ and is

$$E_{\text{MP2}} = 2 \sum_{i,j}^{\text{occ.}} \sum_{a,b}^{\text{vir.}} \langle ij|ab \rangle \langle ab|i j \rangle - \sum_{i,j}^{\text{occ.}} \sum_{a,b}^{\text{vir.}} \langle ij|ab \rangle \langle ab|ji \rangle,$$

(1.2)

where $\langle ij|ab \rangle$ is the two-electron integrals between molecular orbitals (MOs) $ij$ and $ab$ and $\epsilon_p$ is the HF orbital energy of orbital $p$. The motivation of MC-MP2 is to avoid the $O(n^5)$ scaling AO to MO transformation of the two-electron integrals present in the equation Eq. (1.2), by utilizing MC. Instead of evaluating Eq. (1.2), Hylleraas found that one could minimize another expression, now known as the Hylleraas functional

$$E[\Psi^{(1)}] = \langle \Psi^{(1)} | \hat{H}^{(0)} - E^{(0)} | \Psi^{(1)} \rangle + 2 \langle \Psi^{(1)} | \hat{V} | \Psi^{(0)} \rangle,$$

(1.3)

where $\hat{H}^{(0)}$ is the sum of the one-electron Fock operators for each occupied orbital, $E^{(0)}$ is the sum of the HF orbital energies for each occupied orbital, $\hat{V}$ is the perturbation operator and $|\Psi^{(0)}\rangle$ in this case is the Hartree-Fock reference. This Hylleraas functional is variational to the MP2 energy with respect to changes in the MP2 amplitudes. The MP2 amplitudes are then optimized so as to minimize the $E[\Psi^{(1)}]$ with the minimum value of $E[\Psi^{(1)}]$ being equal to $E_{\text{MP2}}$. This is the way the original MP2-R12 energy was evaluated, and how it has been evaluated in our studies as well.
1.3 Original MP2-R12 formulation

1.3.1 Correlation cusp condition

The basis of the explicitly correlated methods is correcting the one-electron basis to have the correct behavior near the correlation cusp, or the cusp of the electronic wavefunction when two electrons coalesce. Kato derived the first form of this cusp condition for general particles. This condition can be stated as

\[
\lim_{r_{12} \to 0} \left( \frac{\partial \Psi}{\partial r_{12}} \right)_{av} = \frac{1}{2} \Psi(r_{12} = 0),
\]

where \( \Psi \) is the many-electron wavefunction, \( r_{12} \) is the inter-electron distance, and \( av \) corresponds to spherical averaging. This demonstrates that in the region of small \( r_{12} \) the wavefunction is linear in \( r_{12} \). This was later integrated by Pack and Byers Brown to define the currently used cusp conditions for singlet and triplet pairs of electrons given as

\[
\Psi = \left[ 1 + \frac{r_{12}}{2(s + 1)} \right] \Phi + O(r_{12}^2),
\]

where \( \Phi \) is the non-interacting component of the wavefunction and \( s = 0 \) and \( 1 \) correspond to singlet and triplet electron pairs respectively.

1.3.2 Klopper and Kutzelnigg

The following analysis is heavily drawn from Klopper and Kutzelnigg’s original R12 papers. The original MP2-R12 formulation by Klopper and Kutzelnigg relies on the previously defined cusp condition to formulate its theory. In order to treat the correlation cusp in practice, we can separate the wavefunction into two different parts

\[
\Psi = g \Phi,
\]

where \( g \) is known as the correlation factor, which accounts for only the correlation cusp, and \( \Phi \) accounts for all other behavior. Kutzelnigg and Klopper found that at \( r_{12} = 0 \), the Coulomb repulsion, \( r_{12}^{-1} \) is singular and the following must be true to cancel the Coulomb singularity

\[
V_{12}g \Phi + [T_{12}, g] \Phi = 0,
\]
where $V_{12}$ is the interaction between the first and second electron and $T_{12}$ is the sum of the kinetic energies of the first and second electron. Evaluating this shows that if $g = \frac{1}{2}r_{12}$

$$V_{12} + [T_{12}, \frac{1}{2}r_{12}]\Phi = \frac{1}{r_{12}} + \frac{1}{2}r_{12}\left(\nabla_1 - \nabla_2\right) = \frac{1}{2r_{12}}(\nabla_1 - \nabla_2) = \hat{U}_{12},$$

(1.8)

where $\vec{r}_{12}$ is the vector $(\vec{r}_1 - \vec{r}_2)$, $\nabla_i$ is the vector operator ($\partial/\partial x_i$, $\partial/\partial y_i$, $\partial/\partial z_i$), and $\hat{U}_{12}$ is Kutzelnigg’s regularization operator. As we can see, satisfying this property of the wavefunction cancels the positive infinite coulomb repulsion with a corresponding negative infinite kinetic energy at the coalescence point $r_{12} = 0$. This can then be done for all electron pairs resulting in the general form

$$g = \prod_{i<j} \left(1 + \frac{1}{2}r_{ij}\right),$$

(1.9)

This guarantees that the first-order cusp conditions are satisfied for all electron pairs.

MP2 is the lowest order non-zero perturbation correlation correction to HF, and thus a natural starting place for the R12 methodology. The R12 correction is added directly onto the first-order wavefunction\textsuperscript{18,19} in Eq. (1.1) giving

$$|\tilde{\Psi}^{(1)}\rangle = \sum_{i<j} t_{ij}^{ab} |ab\rangle + \sum_{i<j} \sum_{k<l} t_{ij}^{kl} \hat{Q}_{12} r_{12} |kl\rangle,$$

(1.10)

where $t_{ij}^{kl}$ are the geminal excitation amplitudes, $r_{12} |kl\rangle$ is the geminal function, and $\hat{Q}_{12}$ is the orthogonality projector. It is defined

$$\hat{Q}_{12} = (1 - \hat{O}_1)(1 - \hat{O}_2) - \hat{V}_1 \hat{V}_2,$$

(1.11)

with

$$\hat{O}_1 = \sum_{\text{occ.}} \varphi_i(r_1) \int dr_3 \varphi^*_i(r_3) \hat{P}_{31},$$

(1.12)

$$\hat{V}_1 = \sum_{\text{vir.}} \varphi_a(r_1) \int dr_3 \varphi^*_a(r_3) \hat{P}_{31},$$

(1.13)

where $\hat{P}_{31}$ permutes all electron coordinates 1 and 3 to the right. The goal of the geminal part of the wavefunction is to correct for the finite-basis error present in the MP2 wavefunction. By inserting pair functions multiplied by $r_{12}$ into the wavefunction, we correct for the behavior around
the correlation cusp where the one-electron basis is fundamentally lacking. The $\hat{Q}_{12}$ operator ensures that the geminal function is orthogonal to the conventional orbital products from HF and MP2. To derive the final R12 corrected energy, this wavefunction is then inserted into the Eq. (1.3), minimizing the energy with respect to the amplitudes. This marked the first step to creating a robust MP2-R12 formalism that would be built on and improved in the later years by many.

### 1.4 Improvements to MP2-R12

Numerous improvements were made to the MP2-R12 formalism from Klopper and Kutzelnigg over the last 30 years. The most relevant will be discussed in the following subsections.

#### 1.4.1 Improved correlation factors

One of the major improvements to the explicitly correlated methodology was the extension to using more complicated correlation factors than linear $r_{12}^{22}$. The function $r_{12}$ does satisfy the cusp condition at $r_{12} = 0$, but it misses out on important intermediate behavior. As stated in Section 1.1, correlation factors used in past implementations were required to be analytically integrable, severely reducing the depth of the analysis. In Chapter 3 we will explore more complicated correlation factors.

#### 1.4.2 Fixed amplitudes

Solving for the geminal amplitudes that minimize the Hylleraas functional comes at a cost of $O(n^6)$, which when applied in MP2 is more expensive than the calculation of the MP2 energy itself. Ten-no proposed the “fixed amplitude ansatz” or the “SP-ansatz” in which the amplitudes are fixed directly to satisfy the cusp conditions by

$$t_{ij}^{mn} = \frac{3}{8} \delta_{mi} \delta_{nj} + \frac{1}{8} \delta_{mj} \delta_{ni}, \quad (1.14)$$

where $\delta_{ij}$ is the kronecker delta between orbitals $ij$. The first term leaves the orbitals the same, and the second term swaps the spatial orbital labels in the pair function. This is because the spatial part of the singlet pair function is symmetric with respect to orbital exchange, whereas the triplet state is antisymmetric. Substituting Eq. (1.14) into Eq. (1.10) results in $3/8 + 1/8 = 1/2$ for the singlet matching the cusp condition from Eq. (1.5) where $s = 0$. Due to the antisymmetry in the triplet
state, the second term is negative leading to $3/8-1/8=1/4$, thereby satisfying the cusp condition when $s = 1$. The cost of using the fixed amplitude ansatz is the introduction of a parameter to account for the system-based properties. Therefore these fixed amplitudes must be paired with a function with an optimizable parameter to derive meaningful results. Ten-no’s choice was a simple Slater-type geminal given by

$$f_{12} = \frac{1 - e^{-\gamma r_{12}}}{\gamma}. \quad (1.15)$$

This has become the most widely used geminal function giving the best overall performance. This new correlation factor pioneered the so-called MP2-F12 methodology, since a function of $r_{12}$ was used instead of just linear $r_{12}$. The error in the MP2-F12 energy is proportional to the square of the error in $\gamma$ leading to a parabolic form of the energy with a clear minimum. For most purposes in organic molecules containing C, H, N, O, and F, choosing $\gamma = 1.1$ is close enough to the minimum to give accurate results.

### 1.5 Metropolis Monte Carlo with importance sampling

The well-known Metropolis MC algorithm is what will be used to integrate the high-dimensional integrals present in the MP2-F12 equations. Importance sampling is a way to sample the space in the regions of highest value of the integrand, or highest importance. The function $f(x)$ to be integrated is decomposed

$$\int f(x)dx = \int \frac{f(x)}{\omega(x)} \omega(x)dx, \quad (1.16)$$

where $\omega(x)$ is the weight function or the importance function. The function $f(x)$ is sampled according to the weight function $\omega(x)$. When sampling directly from $\omega(x)$ is difficult, the Metropolis-Hasting’s algorithm can be used as the rejection sampling technique. How this works in practice is that at each MC step, walkers make a random move in a random direction according to our step size, $\delta x$. The values of the weight function at the original point, $x$, and trial point, $x + \delta x$, are evaluated. A random number, $\alpha$, is then generated on the uniform distribution $[0,1]$. The ratio between the weight functions is then evaluated and the move is accepted or rejected based on

$$x_{n+1} = \begin{cases} x_n + \delta x & \frac{\omega(x_n + \delta x)}{\omega(x_n)} \geq \alpha \\ x_n & \mathrm{otherwise} \end{cases} \quad (1.17)$$
This way if the walker is moving to a point with higher weight or higher importance it will always accept the move while rejecting the moves to lower importance with a probability proportional to the ratio of the weight functions. The accept/rejection ratio can be optimized to a desired value by changing the step size, $\delta x$, with the best performance at a ratio of 50%.

1.6 Redundant walker algorithm

In addition to the conventional Metropolis MC algorithm, the so-called “redundant walker” algorithm developed by Hirata and Willow is used.\textsuperscript{44,55} It is utilized in all methods presented in this thesis. An example 3-electron integral similar to the ones present in this paper is shown below

$$ F = \iiint \frac{f(r_2, r_3)}{r_{12}} dr_1 dr_2 dr_3. \quad (1.18) $$

where $f(r_2, r_3)$ is a non-singular weakly correlating function. Instead of sampling in all 3 dimensions simultaneously, the redundant walker algorithm delegates a set of “pair walkers” to sample $r_1$ and $r_2$, and a different set of “one-electron” walkers to sample $r_3$. This separation is chosen since $r_1$ and $r_2$ are correlated to each other through $r_{12}^{-1}$, and the pair $(r_1, r_2)$ and $r_3$ are only weakly correlated through $f(r_2, r_3)$. They have different weight or importance functions because of the presence of the singularity in the first two coordinates and lack thereof in the third. The weight functions for these two different walkers can be defined as

$$ \omega(r_1, r_2) = \frac{1}{r_{12}}, \quad \omega(r_3) = g(r_3), \quad (1.19) $$

where $g(r_3)$ is a chosen function that emulates the dependence of $f(r_2, r_3)$ on $r_3$. The $m$ pair and one-electron walkers coordinates are generated, and the functions in their integrands and weight functions are evaluated. This process of performing MC moves and recalculating functions is done many times. These are utilized to evaluate the full integral as

$$ F \approx \frac{1}{m^2} \sum_{m_1}^{m} \sum_{m_2}^{m} \frac{f(r_2^{[m_1]}, r_3^{[m_2]})}{r_1^{[m_1]} - r_2^{[m_2]}} \left( \omega(r_1^{[m_1]}, r_2^{[m_1]}) \omega(r_3^{[m_2]}) \right), \quad (1.20) $$

where $r_p^{[m_n]}$ corresponds to the position $r_p$ of walker $m_n$, and the sets $m_1$ and $m_2$ refer to the pair and one-electron walkers respectively. Thus the value of using the redundant walker algorithm is given a small number of redundant walkers, $m$, the cost of propagating the coordinates and evaluating functions thereof is only $O(m)$. While the number of samples generated from all possible pairs is
$O(m^2)$ for an $O(m)$ overall cost benefit. The analysis of the convergence of this integral does not extend to all integrals, but this integral serves as an example of features that are present in nearly all integrals in the MP2-F12 and GF2-F12 equations. Some integrals only have one walker type, and some have one pair walker and three one-electron walkers, making the overall cost benefit difficult to determine.
Chapter 2: MC-MP2-F12-V/VBX method

This chapter is derived from joint work published with S. Hirata, A. E. Doran, J. Zhang, and E. F. Valeev.\textsuperscript{56}

2.1 Introduction

In this work, we propose a solution to all of the shortcomings of MP2-F12 discussed in Section 1.1 by virtue of using a stochastic algorithm, thus presenting a novel MP2-F12 method that can use virtually any geminal, does not require the RI approximation or an ABS, is not predicated upon numerous molecular integrals precomputed or stored on disks, and scales favorably with both system and computer sizes.

It is an extension of the MC-MP2 method,\textsuperscript{44} which is a member of the Brueckner–Goldstone quantum Monte Carlo (BGQMC) family\textsuperscript{57} of methods for both electrons\textsuperscript{43;44;55;58;59;60} and vibrations.\textsuperscript{61;62;63;64} Unlike more conventional quantum Monte Carlo (QMC) methods,\textsuperscript{65;66;67;68;69} MC-MP2 can compute energy differences (such as correlation energies, correlated ionization potentials and electron affinities, and quasiparticle energy bands) directly and not as small differences of noisy total energies. Additionally it does not suffer from any sign problem or fixed-node error and is rigorously size-consistent and thus free from a finite-size error. It is systematically convergent at the exact solutions of the Schrödinger equations as the perturbation rank and basis-set size are increased. Unlike deterministic MP2, MC-MP2 does not need two-electron integrals precomputed or stored either in the AO or MO basis, and is, therefore, more scalable. The operation cost per MC step of MC-MP2 is shown to be linear\textsuperscript{58} with system size and the cost to achieve a given relative statistical uncertainty is found to be cubic.\textsuperscript{70} It can be easily and efficiently parallelized on many central processing units (CPUs)\textsuperscript{55} or on many graphical processing units (GPUs),\textsuperscript{70} sometimes achieving scalability\textsuperscript{70} unprecedented for \textit{ab initio} electron-correlation theories.

In the explicitly correlated extension called MC-MP2-F12 presented here, we exploit another important advantage of MC-MP2, which is its flexibility with various mathematical forms of basis functions. MC-MP2-F12 can thus use virtually any geminal forms as it evaluates necessary integrals numerically by the Metropolis MC method. Furthermore, being a sparse integration method, the MC method’s relative superiority over quadrature grows with dimension; it was argued that the former is more efficient than the latter when the dimension exceeds three.\textsuperscript{71} In this sense, three-
electron (9-dimensional) and higher-dimensional integrals or even two-electron (6-dimensional) integrals may be more suitably handled by MC integrations than by analytical integrations.\textsuperscript{72,73,74,75} The latter are also more expensive to develop and may have a hard ceiling of applicability when the problem size is too large. The former, in contrast, would execute for a far larger problem, only taking longer to converge and give meaningful results.

Also, the MC method makes it unnecessary to factor high-dimensional integrals into lower-dimensional ones in MC-MP2-F12; it can directly evaluate a short sum of high-dimensional integrals, neither requiring the RI approximation and a large ABS nor involving many dense matrix multiplications, which tend not to be scalable. Consequently, the cost of MC-MP2-F12 to achieve a given relative statistical error is found to scale more favorably [$O(n^4)$] than MP2 or MP2-F12 with Ten-no’s fixed amplitudes. Its parallel algorithm of MC-MP2-F12 is also found to exhibit near-perfect scalability up to thousands of processing cores and is essentially free of disk I/O.

Previously Willow et al.\textsuperscript{44} reported a pilot implementation of the MC-MP2-F12 method using the nonvariational V formula (see below) and obtained total correlation energies near their CBS limits using Ten-no’s fixed-amplitude ansatz\textsuperscript{50} with the Slater-type geminal.\textsuperscript{20} Since this (V) formula is not variational with respect to the form of the geminal or the excitation amplitude multiplying its integrals (which are held fixed\textsuperscript{50}), whether it gives reliable relative energies near the CBS limits was not clear. While the cost per MC step for MC-MP2-F12 with the nonvariational (V) formula was shown\textsuperscript{44} to be quadratic with the number of orbitals, the more meaningful cost, i.e., that required to reach a given relative statistical uncertainty, was unknown. Parallel scalability was strongly inferred but never demonstrated, either. Furthermore, the ease of use of any geminal was touted but never exploited. In this paper, we fully develop MC-MP2-F12 using both of the nonvariational (V) and variational (VBX) formulas and answer all of these unanswered questions.

2.2 Theory

2.2.1 $E_{F12}$

In MP2-F12 theory, the MP2 energy and MP2-F12 correction are inherently coupled and must be solved simultaneously.\textsuperscript{76,77} In order to avoid performing the iterations of distinct MC calculations to solve the coupled equations, two main approximations are used to decouple the MP2 and MP2-F12 contributions in the following Monte-Carlo implementation; the Generalized Brillouin
Condition (GBC) and the Extended Brillouin Condition (EBC). They are as follows

\[ \begin{align*}
\hat{F}_n \varphi_i(r_n) &= \epsilon_i \varphi_i(r_n), \\
\hat{F}_n \varphi_a(r_n) &= \epsilon_a \varphi_a(r_n),
\end{align*} \]  

(2.1)

(2.2)

where Eq. 2.1 is the GBC, Eq. 2.2 is the EBC, and \( i \) and \( a \) label an occupied and virtual orbital, respectively. The correlation energy of MP2-F12 theory can then be written as

\[ E_{MP2-F12} = E_{MP2} + E_{F12}, \]  

(2.3)

where \( E_{MP2} \) is the MP2 correlation energy in a finite (AO) basis set and \( E_{F12} \) is the correction for the basis-set-incompleteness error. The latter (the F12 correction) is derived from the Hylleraas functional and consists of three terms:

\[ E_{F12} = 2E_{F12}^V + E_{F12}^B + E_{F12}^X, \]  

(2.4)

with

\[ \begin{align*}
E_{F12}^V &= \sum_{i,j,m,n}^\text{occ.} V_{ij}^{mn}(2t_{ij}^{mn} - t_{ji}^{mn}), \\
E_{F12}^B &= \sum_{i,j,k,l,m,n}^\text{occ.} t_{kl}^{mn} B_{kl}^{ij}(2t_{ij}^{mn} - t_{ji}^{mn}), \\
E_{F12}^X &= - \sum_{i,j,k,l,m,n}^\text{occ.} (\epsilon_m + \epsilon_n) X_{ij}^{kl}(2t_{ij}^{mn} - t_{ji}^{mn}),
\end{align*} \]  

(2.5)

(2.6)

(2.7)

where \( \epsilon_p \) is the \( p \)th Hartree–Fock (HF) orbital energy, \( t_{ij}^{mn} \) is the so-called geminal amplitude, and all summations run over occupied orbitals spanned by the basis set. The other factors are molecular integrals of two electrons, which are written in the standard physicists’ notation as

\[ \begin{align*}
V_{ij}^{mn} &= \langle ij | r_{12}^{-1} \hat{Q}_{12} f_{12} | mn \rangle, \\
B_{mn}^{ij} &= \langle ij | f_{12} \hat{Q}_{12}(\hat{F}_1 + \hat{F}_2) \hat{Q}_{12} f_{12} | mn \rangle, \\
X_{ij}^{kl} &= \langle ij | f_{12} \hat{Q}_{12} f_{12} | mn \rangle,
\end{align*} \]  

(2.8)

(2.9)

(2.10)

where \( f_{12} \) is the geminal (an explicit function of \( r_{12} \)) and \( \hat{F}_n \) is the Fock operator of electron \( n \), i.e.,

\[ \hat{F}_n = \hat{T}_n + \hat{V}_n + \hat{J}_n - \hat{K}_n. \]  

(2.11)
The right-hand side of this equation is the sum of the kinetic-energy operator, the nuclear-attraction operator, the Coulomb operator, and the exchange operator, in this order. Operator \( \hat{Q}_{12} \) is Valeev’s strong-orthogonality projector defined as

\[
\hat{Q}_{12} = (1 - \hat{O}_1)(1 - \hat{O}_2) - \hat{V}_1 \hat{V}_2, \tag{2.12}
\]

with

\[
\hat{O}_1 = \sum_{\text{occ.}} \varphi_i(r_1) \int dr_3 \varphi_i^*(r_3) \hat{P}_{31}, \tag{2.13}
\]

\[
\hat{V}_1 = \sum_{\text{vir.}} \varphi_a(r_1) \int dr_3 \varphi_a^*(r_3) \hat{P}_{31}, \tag{2.14}
\]

where \( \hat{P}_{31} \) permutes all electron coordinates 1 and 3 in what follows. The summation in Eq. (2.14) runs over all virtual orbitals spanned by the basis set.

The F12 correction \( E_{F12} \) is variational with the geminal amplitudes, whose values are, therefore, to be determined by minimizing \( E_{F12} \), through an \( O(n^6) \) process,\textsuperscript{77} where \( n \) is the number of orbitals. For these projectors in \( \hat{Q}_{12} \) conventionally the RI approximation with a large ABS is used in order to keep the integrals in a separable low-dimensional form. The RI approximation is used to higher dimensional electronic integrals into products of 2-electron integrals. However, converging the RI approximation within <0.1% error requires ABS’s with 5d function for H and 5f functions for the heavier atoms.\textsuperscript{49} Since we are using Monte Carlo there is no need for such a large basis set to be used, and the projector is retained as we are evaluating the integral as a single high-dimensional quantity. It can also be shown that at the minimum, \( E_{F12}^B + E_{F12}^X = -E_{F12}^V \) and hence,

\[
E_{F12} = E_{F12}^V. \tag{2.15}
\]

Ten-no showed\textsuperscript{50} that an accurate F12 correction can be obtained by holding the geminal amplitudes fixed at the values dictated by the cusp conditions,\textsuperscript{17,48,80} i.e.,

\[
t_{ij} = \frac{3}{8} \delta_{mi} \delta_{nj} + \frac{1}{8} \delta_{mj} \delta_{ni}, \tag{2.16}
\]

insofar as an appropriate form of the geminal is used, such as Ten-no’s Slater-type geminal,

\[
f_{12} = \frac{1 - e^{-r_{12}}}{\gamma}, \tag{2.17}
\]
where $\gamma$ is an adjustable parameter. With Ten-no's fixed amplitudes, the F12 correction as defined by Eq. (2.4) (the variational $VBX$ formula) is no longer a variational minimum with respect to the “size” ($t_{ij}^{nn}$) of the geminal, but is still variationally bound from below with respect to its “shape” ($\gamma$). The nonvariational ($V$) formula [Eq. (2.15)] is not variational with either. Since the errors in a variational energy have the same sign and tend to also have similar magnitudes, they may cancel with each other to yield accurate energy differences. Below, we quantify the performance of the nonvariational ($V$) and variational ($VBX$) formulas for energy differences. We will also address another important question: what is special about the Slater-type geminal of Eq. (2.17) that makes the fixed-amplitude method work? Below, we explore several geminals that have the right asymptote

$$f_{12} \approx r_{12} \text{ as } r_{12} \to 0.$$  \hspace{1cm} (2.18)

Substituting Eq. (2.16) into Eq. (2.4), we obtain

$$E_{V_{F12}} = \frac{5}{8} \sum_{i,j} V_{ij} - \frac{1}{8} \sum_{i,j} V_{ji},$$  \hspace{1cm} (2.19)

$$E_{B_{F12}} = \frac{7}{32} \sum_{i,j} B_{ij} + \frac{1}{32} \sum_{i,j} B_{ji},$$  \hspace{1cm} (2.20)

$$E_{X_{F12}} = -\frac{7}{32} \sum_{i,j} (\epsilon_i + \epsilon_j)X_{ij} - \frac{1}{32} \sum_{i,j} (\epsilon_j + \epsilon_i)X_{ji}.$$  \hspace{1cm} (2.21)

Before converting them into forms suitable for MC integration, we make use of the GBC and EBC stated earlier,\textsuperscript{21,77} which leads to

$$[\hat{F}_1 + \hat{F}_2, \hat{Q}_{12}] = 0,$$  \hspace{1cm} (2.22)

which, in turn, simplifies the $B$ integrals [Eq. (2.9)] into

$$B_{ij}^j = \langle i| j \rangle f_{12} \hat{Q}_{12}(\hat{F}_1 + \hat{F}_2)Q_{12}f_{12}|i j\rangle$$  \hspace{1cm} (2.23)

$$= \langle i| j \rangle f_{12} \hat{Q}_{12}(\hat{F}_1 + \hat{F}_2)Q_{12}f_{12}|i j\rangle$$

$$+ \langle i| j \rangle f_{12} \hat{Q}_{12} [\hat{F}_1 + \hat{F}_2, \hat{Q}_{12}] f_{12}|i j\rangle$$  \hspace{1cm} (2.24)

$$= \langle i| j \rangle f_{12} \hat{Q}_{12}(\hat{F}_1 + \hat{F}_2)Q_{12}f_{12}|i j\rangle,$$  \hspace{1cm} (2.25)
and

\[ B_{ji}^{ij} = \langle ij | f_{12} \hat{Q}_{12} (\hat{F}_1 + \hat{F}_2) f_{12} | ji \rangle, \]  

(2.26)

where the idempotency of \( \hat{Q}_{12} \) is used, i.e., \( \hat{Q}_{12}^2 = \hat{Q}_{12} \). As will be discussed below, when expanded, each \( \hat{Q}_{12} \) introduces a new electron and increases the dimension of an integral by three [see Eqs. (2.13) and (2.14)]. Therefore, Eq. (2.25) is far more computationally tractable than Eq. (2.23).

Second, using the commutator \( [\hat{F}_1 + \hat{F}_2, f_{12}] \), we rewrite Eq. (2.25) as

\[ B_{ij}^{ji} = \langle ij | f_{12} \hat{Q}_{12} (\hat{F}_1 + \hat{F}_2) f_{12} | ji \rangle \]  

(2.27)

\[ = \langle ij | f_{12} \hat{Q}_{12} [\hat{F}_1 + \hat{F}_2, f_{12}] | ji \rangle + \langle ij | f_{12} \hat{Q}_{12} f_{12} (\hat{F}_1 + \hat{F}_2) | ji \rangle \]  

(2.28)

\[ = \langle ij | f_{12} \hat{Q}_{12} [\hat{F}_1 + \hat{F}_2, f_{12}] | ij \rangle + (\epsilon_i + \epsilon_j) \langle ij | f_{12} \hat{Q}_{12} f_{12} | ij \rangle \]  

(2.29)

\[ = \langle ij | f_{12} \hat{Q}_{12} [\hat{F}_1 + \hat{F}_2, f_{12}] | ij \rangle + (\epsilon_i + \epsilon_j) X_{ij}^{ij}, \]  

(2.30)

and

\[ B_{ji}^{ij} = \langle ij | f_{12} \hat{Q}_{12} [\hat{F}_1 + \hat{F}_2, f_{12}] | ji \rangle + (\epsilon_i + \epsilon_j) X_{ji}^{ij}, \]  

(2.31)

Note that the GBC is invoked in Eq. (2.29). The GBC leads to faster computation time of the B term as it simplifies the integrals into having one projector instead of multiplying a larger 6-dimensional integral or many small-dimensional intermediates in the case of the deterministic method. As we can see from Eq. (2.4), (2.21), and (2.30), using the GBC and the commutator approximation X cancels exactly within B. If the GBC were not invoked, Fock matrix elements between the occupied orbitals would need to be retained and used, and the X term would not vanish within B. The EBC allows for simpler computation due to the ignoring of coupling terms between the MP2 and F12 contributions, in addition to allowing us to utilize the already established MC-MP2 method for a fully stochastic calculation. In addition, the GBC is known to be an accurate approximation, causing only errors on the order of milli Hartrees.\textsuperscript{21,49} The EBC is less accurate, but still tolerable.\textsuperscript{21,49} For a smaller basis set, in fact, it is recommended\textsuperscript{21,49} that the EBC be invoked when the GBC is used.

These two approximations reduce \( E_{F12} \) to

\[ E_{F12} = 2E_{F12}^V + E_{F12}^{BX}, \]  

(2.32)
with

\[ E_{F12}^{BX} = \frac{7}{32} \sum_{i,j} \text{occ. } (BX)_{ij}^{ij} + \frac{1}{32} \sum_{i,j} \text{occ. } (BX)_{ji}^{ij}, \]  

(2.33)

and

\[ (BX)^{ij}_{mn} = \langle ij | f_{12} \hat{Q}_{12} \big[ \hat{F}_{1} + \hat{F}_{2}, f_{12} \big] | mn \rangle. \]  

(2.34)

### 2.2.2 \( E_{F12}^{V} \)

Expanding \( \hat{Q}_{12} \), i.e., substituting Eq. (2.12) into Eq. (2.19), we obtain

\[ E_{F12}^{V} = E_{2e}^{V} + E_{3e}^{V} + E_{4e}^{V}, \]  

(2.35)

with

\[ E_{2e}^{V} = \frac{5}{8} \sum_{i,j} \text{occ. } \langle ij | \frac{f_{12}}{r_{12}} | ij \rangle - \frac{1}{8} \sum_{i,j} \text{occ. } \langle ij | \frac{f_{12}}{r_{12}} | ji \rangle, \]  

(2.36)

\[ E_{3e}^{V} = -\frac{5}{4} \sum_{i,j,k} \text{occ. } \langle ijk | \frac{f_{23}}{r_{12}} | kji \rangle + \frac{1}{4} \sum_{i,j,k} \text{occ. } \langle ijk | \frac{f_{23}}{r_{12}} | kij \rangle, \]  

(2.37)

\[ E_{4e}^{V} = \frac{5}{8} \sum_{i,j,k,l} \text{occ. } \langle ijkl | \frac{f_{34}}{r_{12}} | klji \rangle - \frac{1}{8} \sum_{i,j,k,l} \text{occ. } \langle ijkl | \frac{f_{34}}{r_{12}} | klij \rangle \]  

\[-\frac{5}{8} \sum_{i,j,k,l} \text{vir. } \langle ijab | \frac{f_{34}}{r_{12}} | abi j \rangle + \frac{1}{8} \sum_{i,j,k,l} \text{vir. } \langle ijab | \frac{f_{34}}{r_{12}} | ab ji \rangle, \]  

(2.38)

where the subscripts on the left-hand sides indicate the number of electrons involved and thus the dimension of integrals. They are rewritten in the MC-integrable forms as

\[ E_{2e}^{V} = \int \int dr_{1} dr_{2} F_{2e}^{V}(r_{1}, r_{2}), \]  

(2.39)

\[ E_{3e}^{V} = \int \int \int dr_{1} dr_{2} dr_{3} F_{3e}^{V}(r_{1}, r_{2}, r_{3}), \]  

(2.40)

\[ E_{4e}^{V} = \int \int \int \int dr_{1} dr_{2} dr_{3} dr_{4} F_{4e}^{V}(r_{1}, r_{2}, r_{3}, r_{4}), \]  

(2.41)
with

\[
F_{2e}^V(r_1, r_2) = \frac{5}{8} \frac{f_{12} O_{11} O_{22}}{r_{12}} - \frac{1}{8} \frac{f_{12} O_{12} O_{21}}{r_{12}},
\]

\[
F_{3e}^V(r_1, r_2, r_3) = -\frac{5}{4} \frac{f_{23} O_{13} O_{22} O_{31}}{r_{12}} + \frac{1}{4} \frac{f_{23} O_{12} O_{23} O_{31}}{r_{12}},
\]

\[
F_{4e}^V(r_1, r_2, r_3, r_4) = \frac{5}{8} \frac{f_{34} O_{13} O_{24} O_{31} O_{42}}{r_{12}} - \frac{1}{8} \frac{f_{34} O_{14} O_{23} O_{31} O_{42}}{r_{12}} - \frac{5}{8} \frac{f_{34} O_{13} O_{24} V_{31} V_{42}}{r_{12}} - \frac{1}{8} \frac{f_{34} O_{14} O_{23} V_{31} V_{42}}{r_{12}},
\]

where

\[
O_{pq} = \sum_i \phi_i^* (r_p) \phi_i (r_q),
\]

\[
V_{pq} = \sum_a \phi_a^* (r_p) \phi_a (r_q).
\]

The formalism of this term is unchanged from Ref. 44. Contrasting this with the deterministic method, these integrals for \( E_{F_{12}}^V \) are

\[
E_{Rl}^V = \frac{5}{8} \langle gr \rangle_{ij} - \frac{1}{8} \langle gr \rangle_{ji}
- \frac{5}{4} \langle g \rangle_{ij}^{kp'} \langle r \rangle_{kp'}^{ij} + \frac{1}{4} \langle g \rangle_{ij}^{kp'} \langle r \rangle_{kp'}^{ij}
+ \frac{5}{8} \langle g \rangle_{ij}^{kl} \langle r \rangle_{kl}^{ij} + \frac{1}{8} \langle g \rangle_{ij}^{kl} \langle r \rangle_{kl}^{ij}
- \frac{5}{8} \langle g \rangle_{ij}^{ab} \langle r \rangle_{ab}^{ij} - \frac{1}{8} \langle g \rangle_{ij}^{ab} \langle r \rangle_{ab}^{ij}
\]

(2.47)
where the 3-electron integrals are summed over $p'$, labeling the larger ABS and where

\[
\langle g \rangle_{ij}^{kl} = \langle ij | f_{12} | kl \rangle \tag{2.48}
\]

\[
\langle r \rangle_{ij}^{kl} = \langle ij | r_{12}^{-1} | kl \rangle \tag{2.49}
\]

\[
\langle gr \rangle_{ij}^{kl} = \langle ij | f_{12} r_{12}^{-1} | kl \rangle \tag{2.50}
\]

The RI in this case is only used for the 3-electron integrals and an example of how this is used in practice is

\[
E_{3e}^{(1)} = \langle ij | f_{12} \hat{Q}_{12} r_{12}^{-1} | ij \rangle, \tag{2.51}
\]

\[
E_{3e}^{(1)} = \langle ijk | f_{12} r_{12}^{-1} | kji \rangle, \tag{2.52}
\]

\[
E_{3e}^{(1)} = \langle ij | f_{12} | kp' \rangle \langle kp' | r_{12}^{-1} | ij \rangle, \tag{2.53}
\]

\[
E_{3e}^{(1)} = \langle g \rangle_{ij}^{kp'} \langle r | kj \rangle_{kp'} \tag{2.54}
\]

where the insertion of the RI projector $|\phi_{p'}(r_2)\rangle \langle \phi_{p'}(r_2)|$ is made in between Eq. (2.52) and Eq. (2.53) to separate the 3-dimensional integral into two two-electron integrals.

### 2.2.3 $E_{FI12}^{BX}$

The sum of the $B$ and $X$ terms, Eq. (2.33), is newly considered in this work. Expanding $\hat{Q}_{12}$ in this equation, we find

\[
E_{FI12}^{BX} = E_{2e}^T + E_{3e}^T + E_{4e}^T + E_{3e}^K + E_{4e}^K + E_{5e}^K, \tag{2.55}
\]
with

$$E^{T}_{2e} = \frac{7}{32} \sum_{i,j} \langle ij | f_{12} [\hat{T}_1 + \hat{T}_2, f_{12}] | ij \rangle$$

$$+ \frac{1}{32} \sum_{i,j} \langle ij | f_{12} [\hat{T}_1 + \hat{T}_2, f_{12}] | ji \rangle , \quad (2.56)$$

$$E^{T}_{3e} = -\frac{7}{16} \sum_{i,j,k} \langle ijk | f_{12} [\hat{T}_2 + \hat{T}_3, f_{23}] | kji \rangle$$

$$- \frac{1}{16} \sum_{i,j,k} \langle ijk | f_{12} [\hat{T}_2 + \hat{T}_3, f_{23}] | ki j \rangle , \quad (2.57)$$

$$E^{T}_{4e} = \frac{7}{32} \sum_{i,j,k,l} \langle ijk l | f_{12} [\hat{T}_3 + \hat{T}_4, f_{34}] | kl j \rangle$$

$$+ \frac{1}{32} \sum_{i,j,k,l} \langle ijk l | f_{12} [\hat{T}_3 + \hat{T}_4, f_{34}] | kl j \rangle$$

$$- \frac{7}{32} \sum_{i,j} \sum_{a,b} \langle ijab | f_{12} [\hat{T}_3 + \hat{T}_4, f_{34}] | abij \rangle$$

$$- \frac{1}{32} \sum_{i,j} \sum_{a,b} \langle ijab | f_{12} [\hat{T}_3 + \hat{T}_4, f_{34}] | abji \rangle , \quad (2.58)$$
and

\[ E_{3e}^K = -\frac{7}{32} \sum_{i,j} \langle ij | f_{12} [\hat{\mathbf{K}}_1 + \hat{\mathbf{K}}_2, f_{12}] | ij \rangle \]

\[ -\frac{1}{32} \sum_{i,j} \langle ij | f_{12} [\hat{\mathbf{K}}_1 + \hat{\mathbf{K}}_2, f_{12}] | ji \rangle, \]

(2.59)

\[ E_{4e}^K = +\frac{7}{16} \sum_{i,j,k} \langle ijk | f_{12} [\hat{\mathbf{K}}_2 + \hat{\mathbf{K}}_3, f_{23}] | kji \rangle \]

\[ +\frac{1}{16} \sum_{i,j,k} \langle ijk | f_{12} [\hat{\mathbf{K}}_2 + \hat{\mathbf{K}}_3, f_{23}] | kij \rangle, \]

(2.60)

\[ E_{5e}^K = -\frac{7}{32} \sum_{i,j,k,l} \langle ijk\ell | f_{12} [\hat{\mathbf{K}}_3 + \hat{\mathbf{K}}_4, f_{34}] | ki\ell \rangle \]

\[ -\frac{1}{32} \sum_{i,j,k,l} \langle ijk\ell | f_{12} [\hat{\mathbf{K}}_3 + \hat{\mathbf{K}}_4, f_{34}] | kl\ell \rangle \]

\[ +\frac{7}{32} \sum_{i,j,a,b} \langle ijab | f_{12} [\hat{\mathbf{K}}_3 + \hat{\mathbf{K}}_4, f_{34}] | abij \rangle \]

\[ +\frac{1}{32} \sum_{i,j,a,b} \langle ijab | f_{12} [\hat{\mathbf{K}}_3 + \hat{\mathbf{K}}_4, f_{34}] | abji \rangle, \]

(2.61)

where we use the fact that only the kinetic-energy (\(\hat{T}_n\)) and exchange (\(\hat{\mathbf{K}}_n\)) operators in the Fock operator [Eq. (2.11)] do not commute with \(f_{12}\) and, therefore,

\[ [\hat{F}_1 + \hat{F}_2, f_{12}] = [\hat{\mathbf{V}}_1 + \hat{\mathbf{V}}_2, f_{12}] - [\hat{\mathbf{K}}_1 + \hat{\mathbf{K}}_2, f_{12}]. \]

(2.62)

The subscripts ‘2e,’ ‘3e,’ etc. again refer to the number of electrons that appear in the integrals. It is incremented in the exchange terms [Eqs. (2.59)–(2.61)] because the exchange operator introduces one more electron, just like \(\hat{Q}_{12}\) (see Appendix A.1).

After straightforward, but rather tedious algebra, which is partially computerized in this work,
Table 2.1: Six geminals considered in this work.

<table>
<thead>
<tr>
<th>Geminal</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{12}^{(1)} = (1 - e^{-\gamma r_{12}})/\gamma$</td>
<td>Slater</td>
</tr>
<tr>
<td>$f_{12}^{(2)} = (1 - e^{-\gamma r_{12}})/(\gamma r_{12})$</td>
<td>Gauss</td>
</tr>
<tr>
<td>$f_{12}^{(3)} = \gamma r_{12}/(\gamma + r_{12})$</td>
<td>Rational</td>
</tr>
<tr>
<td>$f_{12}^{(4)} = \ln(1 + \gamma r_{12})/\gamma$</td>
<td>Logarithm</td>
</tr>
<tr>
<td>$f_{12}^{(5)} = \arctan(\gamma r_{12})/\gamma$</td>
<td>Arctangent</td>
</tr>
<tr>
<td>$f_{12}^{(6)} = f_{12}^{(1)}/2 + f_{12}^{(3)}/2$</td>
<td>Hybrid</td>
</tr>
</tbody>
</table>

we arrive at the following MC-integrable expressions:

\[
E_{2e}^T = \int \int d\mathbf{r}_1 d\mathbf{r}_2 F^T_{2e}(\mathbf{r}_1, \mathbf{r}_2) \\
+ \int \int d\mathbf{r}_1 d\mathbf{r}_2 F^T_{2e}(\mathbf{r}_1, \mathbf{r}_2), \quad (2.63)
\]

\[
E_{3e}^T = \int \int \int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 F^T_{3e}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) \\
+ \int \int \int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 F^T_{3e}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3), \quad (2.64)
\]

\[
E_{4e}^T = \int \int \int \int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 d\mathbf{r}_4 F^T_{4e}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) \\
+ \int \int \int \int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 d\mathbf{r}_4 F^T_{4e}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4), \quad (2.65)
\]

\[
E_{5e}^K = \int \int \int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 F^K_{5e}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3), \quad (2.66)
\]

\[
E_{4e}^K = \int \int \int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 d\mathbf{r}_4 F^K_{4e}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4), \quad (2.67)
\]

\[
E_{5e}^K = \int \cdots \int d\mathbf{r}_1 \cdots d\mathbf{r}_5 F^K_{5e}(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4, \mathbf{r}_5). \quad (2.68)
\]
Table 2.2: Components of the commutators with the kinetic-energy operator in Eq. (2.72) for the 6 geminals presented in Table 2.1.

<table>
<thead>
<tr>
<th>Name</th>
<th>( f^{(a)}_{12} )</th>
<th>( f^{(b)}_{12} )</th>
<th>( f^{(c)}_{12} )</th>
<th>( f^{(d)}_{12} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Slater</td>
<td>( -2 e^{-\gamma r_{12}} )</td>
<td>( \gamma e^{-\gamma r_{12}} )</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Gauss</td>
<td>( -2 e^{-\gamma r_{12}^2} )</td>
<td>( 4 \gamma r_{12} e^{-\gamma r_{12}^2} )</td>
<td>( (e^{-\gamma r_{12}^2} - 1)/(\gamma r_{12}^2) + 2 e^{-\gamma r_{12}^2} )</td>
<td>0</td>
</tr>
<tr>
<td>Rational</td>
<td>( -2 \gamma^3/(r_{12} + \gamma)^3 )</td>
<td>0</td>
<td>( \gamma/(\gamma + r_{12}) )</td>
<td>( -\gamma/(\gamma + r_{12})^2 )</td>
</tr>
<tr>
<td>Logarithm</td>
<td>( -2/(1 + \gamma r_{12})^2 )</td>
<td>( -\gamma/(1 + \gamma r_{12})^2 )</td>
<td>( 1/(1 + \gamma r_{12}) )</td>
<td>0</td>
</tr>
<tr>
<td>Arctangent</td>
<td>( -2/(1 + \gamma^2 r_{12}^2)^2 )</td>
<td>0</td>
<td>( 1/(1 + \gamma^2 r_{12}^2) )</td>
<td>0</td>
</tr>
<tr>
<td>Hybrid</td>
<td>( -e^{-\gamma r_{12} - \gamma^2/(r_{12} + \gamma)^3} )</td>
<td>( \gamma e^{-\gamma r_{12}/2} )</td>
<td>( e^{-\gamma r_{12}/2 + \gamma/(2 \gamma + 2 r_{12})} )</td>
<td>( -\gamma/(\gamma + r_{12})^2/2 )</td>
</tr>
</tbody>
</table>

The integrals in the two-electron kinetic-energy contribution [Eq. (2.63)] are given by

\[
\begin{align*}
F_{2e}^{T1}(r_1, r_2) &= \frac{7}{32} f_{12} f^{(a)}_{12} O_{11} O_{22} \\
&+ \frac{7}{32} f_{12} f^{(c)}_{12} (O_{11} O'_{22} - O'_{11} O_{22}) \\
&+ \frac{1}{32} f_{12} f^{(a)}_{12} O_{12} O_{21} \\
&+ \frac{1}{32} f_{12} f^{(c)}_{12} (O_{21} O'_{12} - O'_{21} O_{12}), \\
\end{align*}
\]

(2.69)

\[
\begin{align*}
F_{2e}^{T2}(r_1, r_2) &= \frac{7}{32} f_{12} f^{(b)}_{12} O_{11} O_{22} \\
&+ \frac{7}{32} f_{12} f^{(d)}_{12} (O_{11} O'_{22} - O'_{11} O_{22}) \\
&+ \frac{1}{32} f_{12} f^{(b)}_{12} O_{12} O_{21} \\
&+ \frac{1}{32} f_{12} f^{(d)}_{12} (O_{21} O'_{12} - O'_{21} O_{12}), \\
\end{align*}
\]

(2.70)

with

\[
O'_{pq} = \sum_i^{\text{occ.}} \phi_i^*(r_p) r_{12} \cdot \nabla_q \phi_i(r_q). 
\]

(2.71)

See Appendix A.1 for the derivation of Eqs. (2.69) and (2.70). The integrand of Eq. (2.63) is divided into two terms, \( F_{2e}^{T1} \) and \( F_{2e}^{T2} \), because they have rather different behavior as \( r_{12} \to 0 \), thus requiring different weight functions in MC integrations (see below). The subscripts of \( r_{12} \) in Eq. (2.71) are always ‘12’ and independent of \( p \) or \( q \). This is related to the fact that, in Eqs. (2.69) and (2.70) and throughout our formalism, the singular operator is chosen (by coordinate interchanges)
to be always of the form $r_{12}^{-1}$, so that the same weight function containing $r_{12}^{-1}$ can be consistently applied to these dimensions. This exchange is possible because all orbitals are real. If this was not the case, as is the case in the relativistic version of these methods, such an interchange would not be possible without conjugating the orbitals. The factors of the geminal, $f^{(a)}, f^{(b)}, f^{(c)},$ and $f^{(d)}$, are defined as

\[
\left[ \hat{T}_1 + \hat{T}_2, f_{12} \right] = \frac{f_{12}^{(a)}}{r_{12}} + f_{12}^{(b)} - \left\{ \frac{f_{12}^{(c)}}{r_{12}} + f_{12}^{(d)} \right\} r_{12} \cdot (\nabla_1 - \nabla_2), \tag{2.72}
\]

whose actual forms are compiled in Table 2.2 for the 6 geminals\textsuperscript{81} considered in this work. This form is related to the so-called Kutzelnigg’s regularization operator.\textsuperscript{18} The cusp conditions\textsuperscript{17;48} reduce to

\[
f_{12}^{(a)} \approx -2 \text{ as } r_{12} \to 0, \tag{2.73}
\]

which is satisfied by all the geminals in Table 2.1.

The integrands in the three-electron kinetic contribution [Eq. (2.64)] are given by

\[
F_{3e}^{T_1}(r_1, r_2, r_3) = -\frac{7}{16} f_{12}^{(a)} \frac{O_{31} O_{22} O_{13}}{r_{12}} - \frac{7}{16} f_{12}^{(c)} \frac{(O'_{22} O_{31} - O'_{31} O_{22}) O_{13}}{r_{12}} - \frac{1}{16} f_{12}^{(a)} \frac{O_{32} O_{21} O_{13}}{r_{12}} - \frac{1}{16} f_{12}^{(c)} \frac{(O'_{32} O_{21} - O'_{21} O_{32}) O_{13}}{r_{12}}, \tag{2.74}
\]

\[
F_{3e}^{T_2}(r_1, r_2, r_3) = -\frac{7}{16} f_{12}^{(b)} \frac{O_{31} O_{22} O_{13}}{r_{12}} - \frac{7}{16} f_{12}^{(d)} \frac{(O'_{22} O_{31} - O'_{31} O_{22}) O_{13}}{r_{12}} - \frac{1}{16} f_{12}^{(b)} \frac{O_{32} O_{21} O_{13}}{r_{12}} - \frac{1}{16} f_{12}^{(d)} \frac{(O'_{32} O_{21} - O'_{21} O_{32}) O_{13}}{r_{12}}, \tag{2.75}
\]
whereas those of the four-electron kinetic contribution \([\text{Eq. (2.65)}]\) read

\[
F^{T_1}_{4e}(r_1, r_2, r_3, r_4) = \frac{7}{32} f_{34} f_{12}^{(a)} O_{31} O_{42} O_{13} O_{24} 
+ \frac{7}{32} f_{34} f_{12}^{(c)} (O'_{42} O_{31} - O'_{31} O_{42}) O_{13} O_{24} 
+ \frac{1}{32} f_{34} f_{12}^{(a)} O_{32} O_{41} O_{13} O_{24} 
+ \frac{1}{32} f_{34} f_{12}^{(c)} (O'_{32} O_{41} - O'_{41} O_{32}) O_{13} O_{24} 
- \frac{7}{32} f_{34} f_{12}^{(a)} O_{31} O_{42} V_{13} V_{24} 
- \frac{7}{32} f_{34} f_{12}^{(c)} (O'_{42} O_{31} - O'_{31} O_{42}) V_{13} V_{24} 
- \frac{1}{32} f_{34} f_{12}^{(a)} O_{32} O_{41} V_{13} V_{24} 
- \frac{1}{32} f_{34} f_{12}^{(c)} (O'_{32} O_{41} - O'_{41} O_{32}) V_{13} V_{24},
\]  

and

\[
F^{T_2}_{4e}(r_1, r_2, r_3, r_4) = \frac{7}{32} f_{34} f_{12}^{(b)} O_{31} O_{42} O_{13} O_{24} 
+ \frac{7}{32} f_{34} f_{12}^{(d)} (O'_{42} O_{31} - O'_{31} O_{42}) O_{13} O_{24} 
+ \frac{1}{32} f_{34} f_{12}^{(b)} O_{32} O_{41} O_{13} O_{24} 
+ \frac{1}{32} f_{34} f_{12}^{(d)} (O'_{32} O_{41} - O'_{41} O_{32}) O_{13} O_{24} 
- \frac{7}{32} f_{34} f_{12}^{(b)} O_{31} O_{42} V_{13} V_{24} 
- \frac{7}{32} f_{34} f_{12}^{(d)} (O'_{42} O_{31} - O'_{31} O_{42}) V_{13} V_{24} 
- \frac{1}{32} f_{34} f_{12}^{(b)} O_{32} O_{41} V_{13} V_{24} 
- \frac{1}{32} f_{34} f_{12}^{(d)} (O'_{32} O_{41} - O'_{41} O_{32}) V_{13} V_{24}.
\]  

(2.76)
The integrands in the exchange contributions [Eqs. (2.66)–(2.68)] are found to be

\[
F^K_{3e}(r_1, r_2, r_3) = \frac{7}{16} \frac{f_{23}(f_{23} - f_{13})}{r_{12}} O_{12} O_{33} O_{21} + \frac{1}{16} \frac{f_{23}(f_{23} - f_{13})}{r_{12}} O_{13} O_{32} O_{21}, \quad (2.78)
\]

\[
F^K_{4e}(r_1, r_2, r_3, r_4) = -\frac{7}{16} \frac{f_{24}(f_{23} - f_{13})}{r_{12}} O_{43} O_{21} O_{34} O_{12} - \frac{1}{16} \frac{f_{24}(f_{23} - f_{13})}{r_{12}} O_{41} O_{23} O_{34} O_{12} - \frac{7}{16} \frac{f_{34}(f_{14} - f_{24})}{r_{12}} O_{32} O_{44} O_{13} O_{21} - \frac{1}{16} \frac{f_{34}(f_{14} - f_{24})}{r_{12}} O_{34} O_{42} O_{13} O_{21}, \quad (2.79)
\]

and

\[
F^K_{5e}(r_1, r_2, r_3, r_4, r_5) = \frac{7}{16} \frac{f_{35}(f_{14} - f_{24})}{r_{12}} O_{32} O_{54} O_{31} O_{45} O_{21} + \frac{1}{16} \frac{f_{35}(f_{14} - f_{24})}{r_{12}} O_{34} O_{52} O_{31} O_{45} O_{21} - \frac{7}{16} \frac{f_{35}(f_{14} - f_{24})}{r_{12}} O_{32} O_{54} V_{31} V_{45} O_{21} - \frac{1}{16} \frac{f_{35}(f_{14} - f_{24})}{r_{12}} O_{34} O_{52} V_{31} V_{45} O_{21}. \quad (2.80)
\]

See Appendix A.1 for the derivation of Eq. (2.78) as an example.

### 2.3 Monte Carlo algorithm

Each of the contributions to the F12 correction is evaluated by the Metropolis MC method with the redundant-walker convergence-acceleration technique.\textsuperscript{55} Identifying and using an appropriate weight function is essential for the viability (let alone efficiency) of any MC integration. A weight function should be positive everywhere, be analytically integrable, have the same singularity as the integrand, and generally behave like the integrand. The closer the weight function is to the true integrand, the better convergence properties the MC calculation will have.
2.3.1 $E_{V12}^V$

The $V$ term [Eq. (2.35)] is evaluated as

$$E_{2e}^V = \int\int dr_1 dr_2 F_{2e}^V(r_1, r_2) \approx \frac{1}{N} \sum_{n=1}^{N} \frac{F_{2e}^V(r_1^{[n]}, r_2^{[n]})}{w_{2e}(r_1^{[n]}, r_2^{[n]})},$$

(2.81)

$$E_{3e}^V = \int\int\int dr_1 dr_2 dr_3 F_{3e}^V(r_1, r_2, r_3) \approx \frac{1}{N} \sum_{n=1}^{N} \frac{1}{m} \sum_{k=1}^{m} \frac{F_{3e}^V(r_1^{[n]}, r_2^{[n]}, r_3^{[n]})}{w_{2e}(r_1^{[n]}, r_2^{[n]})w_{1e}(r_3^{[n]})},$$

(2.82)

and

$$E_{4e}^V = \int\int\int\int dr_1 dr_2 dr_3 dr_4 F_{4e}^V(r_1, r_2, r_3, r_4) \approx \frac{1}{N} \sum_{n=1}^{N} \frac{2!}{m(m-1)} \sum_{l=k+1}^{m} \sum_{l=1}^{m} \frac{F_{4e}^V(r_1^{[n]}, r_2^{[n]}, r_3^{[n]}, r_4^{[n]})}{w_{2e}(r_1^{[n]}, r_2^{[n]})w_{1e}(r_3^{[n]})w_{1e}(r_4^{[n]})},$$

(2.83)

where $N$ is the total number of MC steps and $m$ is the number of “redundant walkers” (see below for the definition).

Electron pairs with coordinates $\{r_1^{[n]}, r_2^{[n]} | 1 \leq n \leq N\}$ are distributed randomly but according to the normalized two-electron weight function of the form,

$$w_{2e}(r_1, r_2) = \frac{1}{N_{2e}} \frac{g(r_1)g(r_2)}{r_{12}},$$

(2.84)

where $g(r)$, in our implementation, is chosen to be a sum of two atom-centered $s$-type GTOs per atom,

$$g(r) = \sum_{A=1}^{n_{atom}} \left( c_A^{(1)} e^{-\xi_A^1 |r-r_A|^2} + c_A^{(2)} e^{-\xi_A^2 |r-r_A|^2} \right).$$

(2.85)

Here, $n_{atom}$ is the number of atoms and $r_A$ is the position of the $A$th atom. The normalization coefficient, $N_{2e}$, can be evaluated analytically. Although the $r_{12}^{-1}$ singularity in $F_{2e}^V$ [Eq. (2.42)] is analytically removed by $f_{12}$ in its numerator, we elect to use the above weight function that makes
MC algorithm sample more heavily at short $r_{12}$ distances. This is appropriate or even necessary because $F_{V}^{2e}$ is expected to vary more rapidly at short $r_{12}$ distances, even after the singularity is removed.\textsuperscript{83,84,85} In $F_{3e}^{V}$ [Eq. (2.43)] and $F_{4e}^{V}$ [Eq. (2.44)], the $r_{12}^{-1}$ singularity remains and the use of the weight function that cancels it is essential. A random distribution according to $w_{2e}$ is achieved by the Metropolis algorithm.\textsuperscript{55,86} Originally, the $g(r)$ were chosen to be the HF density, but this turned out to be too tight of a distribution causing the quotient in Eq. (2.84) to become very large due to the diffuseness of virtual orbitals present in the integrals.

A distribution of one-electron coordinates \{ $r_{3k}$ $|$ $1 \leq n \leq N$ \} is generated randomly but according to the normalized one-electron weight function,

$$ w_{1e}(r_3) = \frac{g(r_3)}{N_{1e}}, \tag{2.86} $$

where $N_{1e}$ is again analytically determined.\textsuperscript{82} Electron 3 or 4 is not strongly coupled with the others in integrand $F_{3e}^{V}$ or $F_{4e}^{V}$, and, therefore, one-electron “walkers” whose distributions resemble the molecule’s electron density are appropriate in these dimensions. In the redundant-walker algorithm,\textsuperscript{44,55} $m$ such independent distributions (redundant walkers) are generated ($1 \leq k \leq m$), so that each of these $m$ one-electron walkers can be used in Eq. (2.82) and $m(m-1)/2$ distinct pairs of one-electron walkers in Eq. (2.83). Since generating $m$ distributions increases the cost by no more than a factor of $m$, but it increases the number of distinct summands (samples) by $O(m^2)$, this algorithm increases the sampling efficiency of Eq. (2.83) by $O(m)$. However, Eq. (2.81) is unaffected by the algorithm. Therefore, the overall performance boost by the redundant-walker algorithm in MC-MP2-F12 is hard to predict. The $V$ contribution implemented in this manner was reported in Ref.\textsuperscript{44}.

### 2.3.2 $E_{F12}^{BX}$

The sum of the $B$ and $X$ terms [Eq. (2.55)] consists of two-, three-, and four-electron kinetic-energy contributions as well as three-, four-, and five-electron exchange contributions, which are as high
as 15-dimensional. They are evaluated as

\[
E_{2e}^T \approx \frac{1}{N} \sum_{n=1}^{N} \frac{F_{T1}(r_1^{[n]}, r_2^{[n]})}{w_{2e}(r_1^{[n]}, r_2^{[n]})} + \frac{1}{N} \sum_{n=1}^{N} \frac{2!}{m(m-1)} \sum_{k=1}^{m} \sum_{l=k+1}^{m} \frac{F_{T2}(r_{1k}^{[n]}, r_{2l}^{[n]})}{w_{1e}(r_{1k}^{[n]})w_{1e}(r_{2l}^{[n]})},
\]

(2.87)

\[
E_{3e}^T \approx \frac{1}{N} \sum_{n=1}^{N} \frac{1}{m} \sum_{k=1}^{m} \frac{F_{T1}(r_1^{[n]}, r_2^{[n]}, r_3^{[n]})}{w_{2e}(r_1^{[n]}, r_2^{[n]})w_{1e}(r_3^{[n]})} + \frac{1}{N} \sum_{n=1}^{N} \frac{3!}{m(m-1)(m-2)} \times \sum_{k=1}^{m-2} \sum_{l=k+1}^{m-1} \sum_{h=l+1}^{m} \frac{F_{T2}(r_{1k}^{[n]}, r_{2l}^{[n]}, r_{3h}^{[n]})}{w_{1e}(r_{1k}^{[n]})w_{1e}(r_{2l}^{[n]})w_{1e}(r_{3h}^{[n]})},
\]

(2.88)

and

\[
E_{4e}^T \approx \frac{1}{N} \sum_{n=1}^{N} \frac{2!}{m(m-1)} \times \sum_{k=1}^{m-1} \sum_{l=k+1}^{m} \frac{F_{T1}(r_1^{[n]}, r_2^{[n]}, r_3^{[n]}, r_4^{[n]})}{w_{2e}(r_1^{[n]}, r_2^{[n]})w_{1e}(r_3^{[n]})w_{1e}(r_4^{[n]})} + \frac{1}{N} \sum_{n=1}^{N} \frac{4!}{m(m-1)(m-2)(m-3)} \sum_{k=1}^{m-3} \sum_{l=k+1}^{m-2} \sum_{h=l+1}^{m-1} \sum_{i=h+1}^{m} \frac{F_{T2}(r_{1k}^{[n]}, r_{2l}^{[n]}, r_{3h}^{[n]}, r_{4i}^{[n]})}{w_{1e}(r_{1k}^{[n]})w_{1e}(r_{2l}^{[n]})w_{1e}(r_{3h}^{[n]})w_{1e}(r_{4i}^{[n]})},
\]

(2.89)

where \(w_{2e}\) is the weight function for an electron-pair walker for strongly coupled integration variables \((r_1, r_2)\), while \(w_{1e}\) is the weight function for \(m\) independent one-electron walkers for relatively uncoupled variables. These weight functions are identical to Eqs. (2.84) and (2.86), respectively, and hence the same electron-pair and one-electron walkers for the \(V\) term can be reused for the \(BX\) term (as well as for MC-MP2). The sampling efficiency of the first term of Eq. (2.87) is unchanged by the redundant-walker algorithm and that of the second term is increased by a factor of \(O(m)\). The \(m\)-dependence of the number of samplings in individual terms in Eqs. (2.88) and (2.89) can be inferred similarly, but that of the overall performance is again hard to predict, but is expected to be small.
The exchange contributions are evaluated as

$$E^K_{3e} \approx \frac{1}{N} \sum_{n=1}^{N} \frac{1}{m} \sum_{k=1}^{m} \frac{F^K_{3e}(r_1^{[n]}, r_2^{[n]}, r_3^{[n]}_k)}{w_2e(r_1^{[n]}, r_2^{[n]})w_1e(r_3^{[n]}_k)},$$

(2.90)

$$E^K_{4e} \approx \frac{1}{N} \sum_{n=1}^{N} \frac{2!}{m(m-1)} \sum_{k=1}^{m-1} \sum_{l=k+1}^{m} \frac{F^K_{4e}(r_1^{[n]}, r_2^{[n]}, r_3^{[n]}_k, r_4^{[n]}_l)}{w_2e(r_1^{[n]}, r_2^{[n]})w_1e(r_3^{[n]}_k)w_1e(r_4^{[n]}_l)},$$

(2.91)

and

$$E^K_{5e} \approx \frac{1}{N} \sum_{n=1}^{N} \frac{3!}{m(m-1)(m-2)} \sum_{k=1}^{m-2} \sum_{l=k+1}^{m-1} \sum_{h=l+1}^{m} \frac{F^K_{5e}(r_1^{[n]}, r_2^{[n]}, r_3^{[n]}_k, r_4^{[n]}_l, r_5^{[n]}_h)}{w_2e(r_1^{[n]}, r_2^{[n]})w_1e(r_3^{[n]}_k)w_1e(r_4^{[n]}_l)w_1e(r_5^{[n]}_h)},$$

(2.92)

again reusing the same electron-pair walkers and \(m\) independent one-electron walkers. The use of \(w_2e\) for variables \(r_1\) and \(r_2\) and of \(w_1e\) for the other variables can be rationalized by the structure (occurrence of singularity) in the integrands [Eqs. (2.78)–(2.80)].

### 2.3.3 Statistical uncertainty

The MC integrals for both the nonvariational \((V)\) [Eq. (2.15)] and variational \((V BX)\) [Eq. (2.32)] formulas can be written in a unified form as

$$E_{F12} = \lim_{N \to \infty} I_N,$$

(2.93)

$$I_N = \frac{1}{N} \sum_{n=1}^{N} I^{[n]},$$

(2.94)

where \(I^{[n]}\) collects all summands of the outermost summations in Eqs. (2.81)–(2.83), (2.87)–(2.89), and/or (2.90)–(2.92). The statistical uncertainty \(\sigma_N\) in \(I_N\) at the \(N\)th MC step can be estimated as

$$\sigma^2_N = \frac{1}{N^2} \sum_{n=1}^{N} \left(I^{[n]} - I_N\right)^2.$$

(2.95)
However, this estimate is well known\textsuperscript{87} to be an underestimation of the true statistical uncertainty because the coordinates of a walker are correlated across several MC steps. A more accurate estimate is obtained by the blocking algorithm of Flyvbjerg and Petersen\textsuperscript{87} as

\[
\sigma_N^2 = \frac{N_b^2}{N^2} \sum_{n=1}^{N/N_b} \left\{ \frac{1}{N_b} \sum_{n'=1}^{N_b} I^{(n-1)N_b+n'} - I_N \right\}^2,
\]

(2.96)

where \(N_b\) is the block size, which is to be gradually enlarged until \(\sigma_N\) plateaus. The relative error at the \(N\)th MC step is, therefore,

\[
\sigma_{\text{rel}} = \frac{\sigma_N}{I_N}.
\]

(2.97)

The error in MC-MP2-F12 (relative to deterministic MP2-F12) seems to be the statistical uncertainty only, and there is no bias (i.e., systematic error).\textsuperscript{88} See Sec. 2.4.3 for more on this issue.

### 2.4 Results and discussion

A massively parallel MC-MP2-F12 program was implemented using the redundant-walker algorithm. Both nonvariational \(V\) formula\textsuperscript{44} [Eq. (2.15)] and variational \(V BX\) formula [Eq. (2.32)] were used, and the notational distinction is made in this article in the parentheses following the method label as in MC-MP2-F12(\(V\)) and MC-MP2-F12(\(V BX\)).

The number of redundant walkers (\(m\) in various equations in Sec. 2.3) was 40 and the block size [\(N_b\) in Eq. (2.96)] was 6. The frozen-core approximation was used in all cases. A F12 correction thus obtained was added to the corresponding MP2 energy obtained by the conventional deterministic algorithm in \texttt{nwchem}.\textsuperscript{89} This is merely to isolate the performance characteristics of MC-MP2-F12 from those of MC-MP2 because the result of the latter becomes immediately available from the former with no extra cost.

The CBS limits of the MP2 energy were extrapolated applying the \(X^{-3}\) formula\textsuperscript{90} to the deterministic values obtained with the aug-cc-pVQZ and aug-cc-pV5Z basis sets. Hereafter, we abbreviate the cc-pVXZ basis set as “XZ” and the aug-cc-pVXZ basis set as “AXZ,” where \(X = D, T, Q,\) or 5.
2.4.1 Geminals and V versus VBX formulas

The MP2 correlation energies for H$_2$O and CH$_4$ at the geometries of Bak et al.\textsuperscript{91} were computed with the MC-MP2-F12 method with the DZ and TZ basis sets. Six geminals\textsuperscript{81} listed in Table 2.2 were used (whose forms near $r_{12} = 0$ are visualized in Fig. 2.1). Recall that any differentiable function can be employed as a geminal in MC-MP2-F12(VBX) or any function (even a numerically defined one) in MC-MP2-F12(V). This is one of the unique advantages of MC-MP2-F12, shared by QMC, the latter using the Jastrow factor routinely.\textsuperscript{92,93,94,95} See also Nooijen and Bartlett,\textsuperscript{81} who proposed many geminals including those used here, Monkhorst\textsuperscript{96} for a rational geminal whose form is dictated by its equation of motion in the first order and asymptotic regions, Grüneis et al.,\textsuperscript{97} who used a “Yukawa–Coulomb” geminal with their planewave MP2-F12, and Silkowski et al.\textsuperscript{98} for a range-separated geminal.

Figure 2.2 shows the $\gamma$-dependence of the MC-MP2-F12 energy of H$_2$O. The nonvariational V data are plotted as green curves, while the variational VBX ones as purple curves. The uncorrected MP2 energies with the DZ and TZ basis sets as well as the CBS limit are indicated as red lines. Figure 2.3 plots the same for CH$_4$.

First, focusing on the bounded VBX curves (purple), we find that all geminals work reasonably
Figure 2.2: The MP2 correlation energies of H$_2$O as a function of $\gamma$ of the 6 geminals listed in Table 2.1. The geometry was taken from Ref. 91 and the CBS limit was obtained by the $X^{-3}$ extrapolation using the AQZ and A5Z basis sets. The number of MC steps was $8.65 \times 10^7$.

well with the exception of geminal 2; the curves from all geminals except 2 are not strongly dependent on $\gamma$ and close to the CBS limit in a wide range of $\gamma$. Geminal 1 is the Slater-type geminal introduced by Ten-no$^{20,99}$ and is widely regarded as one of the best-performing geminals. Geminal 2 is the Gauss-type geminal, which should be distinguished from the Gaussian geminal that does not satisfy the cusp conditions.$^{20}$ Its performance is rather poor as its MC-MP2-F12(VBX)/DZ energy falls short of the MP2/TZ energy at any value of $\gamma$.

Comparing these performance data with Fig. 2.1, we notice that only geminal 2 has a rather different long-range behavior than the rest. Unlike the other geminals, which rise monotonically, geminal 2 increases and then turns to decrease with $r_{12}$, meaning that its correlation hole may be 2$s$-orbital-like, when the true correlation hole is 1$s$-orbital-like.$^{85}$ It has been shown$^{22,51}$ that the
Figure 2.3: The MP2 correlation energies of CH₄ as a function of γ of the 6 geminals listed in Table 2.1. The geometry was taken from Ref.⁹¹ and the CBS limit was obtained by the X⁻³ extrapolation using the AQZ and A5Z basis sets. The number of MC steps was 8.65 × 10⁷.

The $r_{12}^2$ component in the Taylor-series expansion of a geminal plays an important role in describing the correlation hole correctly, but among the 6 geminals studied here, only geminals 2 and 5 lack this component (the VBX curves of geminal 5 are equally unstable with γ as geminal 2, but as will be shown below the relative energies from geminal 5 seem reasonable). These may explain the poor performance of geminal 2.

Next, turning our attention to the nonvariational V curves (green), we observe that they strongly depend on γ and are usually unbounded. For instance, using geminal 5 with γ < 1.2 a.u., the F12 corrections overshoot the correct values in both H₂O and CH₄, rendering the MC-MP2-F12(V)/DZ energies more negative than the MP2/CBS values, a clear violation of Hylleraas’ variational principle. This may or may not become troublesome in chemical applications of MP2-F12, wherein
relative energies are often sought; if the errors from the CBS limits have opposite signs let alone
greatly different magnitudes across molecules, the errors in relative energies can be amplified, pos-
sibly undoing improvements made by F12 in the total correlation energies. In Sec. 2.4.3, we will
examine this numerically.

As will be shown in Sec. 2.4.4, the nonvariational V formula has the advantage over the vari-
ational VBX formula of being 2–10 times faster. Can one then use the V formula and reliably
determine the CBS limits? We have conceived of two ways to do this. The first is to combine
two geminals with opposite γ-dependences, such that the resulting geminal is nearly independent
of γ. Geminal 6 is indeed constructed in this way as an average of geminals 1 and 3, whose V
curves have opposite γ-dependences. Consequently, the MC-MP2-F12(V) energies of geminal 6
are stable with variation in γ and also much closer to the CBS limits than those from geminal 1
or 3 alone at any γ. We will examine if this construction improves total correlation energies for a
wider array of molecules in Sec. 2.4.2 and relative energies in Sec. 2.4.3.

The second method attempts to locate the value of γ at which the nonvariational V formula
yields the CBS limit. It does so by combining the results of two basis sets. At a given basis set, the
VBX energy is closest to the CBS limit at γ_min that minimizes it:

$$E_{\text{MP2/CBS}} \approx E_{\text{MP2/XZ}} + 2E_{\text{F12}(\gamma_{\text{min}})/XZ} + E_{\text{BX}}(\gamma_{\text{min}})/XZ.$$ \hspace{1cm} (2.98)

Note the approximate equality in the above expression, which is because the left-hand side is a
lower bound of the right-hand side. At the minimum, however, the right-hand side should be
nearly equal to the V energy as per Eq. (3.7) insofar as Ten-no’s fixed amplitudes approximate
well the true variationally optimized geminal amplitudes:

$$2E_{\text{F12}(\gamma_{\text{min}})/XZ} + E_{\text{BX}}(\gamma_{\text{min}})/XZ = E_{\text{F12}(\gamma_{\text{min}})/XZ}.$$ \hspace{1cm} (2.99)

This seems, in fact, borne out in most of the plots of Figs. 2.2 and 2.3; the intersection of the VBX
and V curves of the same basis set occurs at the minimum of the former (except for the curves of
geminal 2). This once again attests to the excellent transferability of ’s fixed amplitudes across all
geminals that have a physically reasonable functional form.

If we now assume the equality in Eq. (2.98) (instead of the approximate equality), we have

$$E_{\text{MP2/CBS}} = E_{\text{MP2/DZ}} + E_{\text{F12}(\gamma_{\text{min}})/DZ} \hspace{1cm} (2.100)$$

$$= E_{\text{MP2/TZ}} + E_{\text{F12}(\gamma_{\text{min}})/TZ}.$$ \hspace{1cm} (2.101)
suggesting that $\gamma_{\text{min}}$ is at the intersection of the $V$ curves of two different basis sets, at which the energy is the CBS limit. Applying this method to the linearly interpolated $V$ curves of geminal 1, 3, and 4, we recover respectively 100.2%, 99.7%, and 99.2% of the CBS limit of H$_2$O. They are closer to 100% than 94.8% or 98.0% recuperated, respectively, by the $VBX$ method with the DZ and TZ basis sets. Likewise, this intersection method recovers 100.4%, 100.1%, and 100.1% of the CBS limit for CH$_4$ using geminals 1, 3, and 4, respectively. Notice how closely the two green curves (of geminals 1, 3, and 4) and the MP2/CBS line (red) meet at one and the same point for both H$_2$O and CH$_4$. Although it is at present unclear whether the efficacy of the intersection method occurs universally across many molecules, we have a potentially practical way of capturing nearly 100% of the CBS limits inexpensively using only the nonvariational $V$ formula.

### 2.4.2 Correlation energies

The F12 corrections and statistical uncertainties were calculated for 17 molecules at the geometries of Bak et al.\textsuperscript{91} by MC-MP2-F12 with the ADZ and ATZ basis set using the $V$ or $VBX$ formula. The MP2 energies were also computed for the same molecules with the AXZ basis set ($X = D, T, Q,$ and 5) using nwchem.\textsuperscript{89}

Table 2.3 compiles the MP2 correlation energies in the increasing order of their magnitude and Table 2.4 compiles the MC-MP2-F12 correlation corrections. The MC-MP2-F12 calculations used the Slater-type geminal with $\gamma = 1.1$ a.u. With the ADZ basis set, the F12 correction brings an average error from the CBS limit from 107.1 m$E_h$ to 6.6 m$E_h$ (with the $V$ formula) or 6.9 m$E_h$ ($VBX$). The latter two errors are smaller than the average error of 9.9 m$E_h$ in MP2/aug-cc-pV5Z. With the ATZ basis set, MP2 has an average error of 42.5 m$E_h$, whereas MC-MP2-F12($V$) has only 3.0 m$E_h$ and MC-MP2-F12($VBX$) 3.5 m$E_h$. The statistical uncertainties are on the order of a few tenths of 1 m$E_h$ with the largest being 1.2 m$E_h$ after $1.44 \times 10^8$ MC steps using 40 redundant walkers. They can furthermore be arbitrarily (albeit slowly) compressed by running a longer MC run.

Figure 2.4 plots the proportion of the CBS limits recovered by MP2 and MC-MP2-F12($VBX$). MP2 recovers on average 74.6, 90.2, 95.6, and 97.7% of the CBS limits with the ADZ, ATZ, AQZ, and A5Z basis sets, respectively. MC-MP2-F12($VBX$)/ADZ captures 98.5% of the CBS limits, which is already greater than that of MP2/A5Z. The ratio goes up further to 99.3% with MC-MP2-F12($VBX$)/ATZ.

Therefore, MC-MP2-F12 works exceedingly well for the total correlation energies, with either the $V$ or $VBX$ formula. Comparing $V$ and $VBX$, the former performs slightly better than the latter. This is not surprising because errors from a variational formula are always positive and cannot
be accidentally small, while the results from a nonvariational formula can scatter in both higher or lower sides of the correct values and can accidentally agree more accurately with the latter. As shown in the next subsection, VBX becomes more accurate than V for the relative energies because systematic cancellation of positive errors occurs in the former (but not in the latter) when energy differences are taken.

In Table 2.5, we revisit the question of the relative performance of geminals for total correlation energies. With the ADZ basis set, \(^{1}\)'s Slater-type geminal (geminal 1) performs the best by a large margin especially in the V formula with an average error of 6.6 \(mE_h\). The hybrid geminal (geminal 6), designed to be insensitive to \(\gamma\) in the V formula, comes next in performance with an average error of 12.8 \(mE_h\). Geminal 2 is particularly poorly performing. Switching from the nonvariational (V) to variational (VBX) formula, the average errors of some geminals decrease, while those of the others increase, reflecting the fact that the values of \(\gamma\) are not optimal for either formula. With the ATZ basis set, the hybrid geminal (geminal 6) becomes the most accurate in both the V and VBX formulas.

<table>
<thead>
<tr>
<th>Molecule</th>
<th>ADZ</th>
<th>ATZ</th>
<th>AQZ</th>
<th>A5Z</th>
<th>CBS(^1)</th>
</tr>
</thead>
<tbody>
<tr>
<td>H(_2)</td>
<td>-0.0273</td>
<td>-0.0320</td>
<td>-0.0333</td>
<td>-0.0337</td>
<td>-0.0343</td>
</tr>
<tr>
<td>CH(_2)</td>
<td>-0.1154</td>
<td>-0.1410</td>
<td>-0.1493</td>
<td>-0.1524</td>
<td>-0.1557</td>
</tr>
<tr>
<td>CH(_4)</td>
<td>-0.1677</td>
<td>-0.2008</td>
<td>-0.2111</td>
<td>-0.2149</td>
<td>-0.2188</td>
</tr>
<tr>
<td>NH(_3)</td>
<td>-0.1992</td>
<td>-0.2401</td>
<td>-0.2537</td>
<td>-0.2588</td>
<td>-0.2643</td>
</tr>
<tr>
<td>H(_2)O</td>
<td>-0.2193</td>
<td>-0.2683</td>
<td>-0.2859</td>
<td>-0.2929</td>
<td>-0.3002</td>
</tr>
<tr>
<td>C(_2)H(_2)</td>
<td>-0.2223</td>
<td>-0.2798</td>
<td>-0.3012</td>
<td>-0.3100</td>
<td>-0.3192</td>
</tr>
<tr>
<td>C(_2)H(_4)</td>
<td>-0.2628</td>
<td>-0.3137</td>
<td>-0.3311</td>
<td>-0.3376</td>
<td>-0.3445</td>
</tr>
<tr>
<td>HF</td>
<td>-0.2847</td>
<td>-0.3399</td>
<td>-0.3583</td>
<td>-0.3651</td>
<td>-0.3722</td>
</tr>
<tr>
<td>HNC</td>
<td>-0.2824</td>
<td>-0.3377</td>
<td>-0.3573</td>
<td>-0.3650</td>
<td>-0.3731</td>
</tr>
<tr>
<td>HCN</td>
<td>-0.2943</td>
<td>-0.3502</td>
<td>-0.3699</td>
<td>-0.3777</td>
<td>-0.3859</td>
</tr>
<tr>
<td>CO</td>
<td>-0.2992</td>
<td>-0.3607</td>
<td>-0.3837</td>
<td>-0.3932</td>
<td>-0.4031</td>
</tr>
<tr>
<td>N(_2)</td>
<td>-0.3173</td>
<td>-0.3796</td>
<td>-0.4018</td>
<td>-0.4109</td>
<td>-0.4203</td>
</tr>
<tr>
<td>CH(_2)O</td>
<td>-0.3327</td>
<td>-0.4021</td>
<td>-0.4271</td>
<td>-0.4370</td>
<td>-0.4474</td>
</tr>
<tr>
<td>HNO</td>
<td>-0.3686</td>
<td>-0.4455</td>
<td>-0.4729</td>
<td>-0.4841</td>
<td>-0.4957</td>
</tr>
<tr>
<td>H(_2)O(_2)</td>
<td>-0.4154</td>
<td>-0.5079</td>
<td>-0.5411</td>
<td>-0.5545</td>
<td>-0.5685</td>
</tr>
<tr>
<td>F(_2)</td>
<td>-0.4280</td>
<td>-0.5361</td>
<td>-0.5758</td>
<td>-0.5927</td>
<td>-0.6105</td>
</tr>
<tr>
<td>CO(_2)</td>
<td>-0.5053</td>
<td>-0.6117</td>
<td>-0.6514</td>
<td>-0.6676</td>
<td>-0.6846</td>
</tr>
</tbody>
</table>

Error\(^2\) | 0.1071 | 0.0425 | 0.0193 | 0.0099 | 0.0000 |

\(^1\) Geometries were taken from Ref.\(^9\).
\(^2\) The X-\(^3\) extrapolation using the MP2/AQZ and MP2/A5Z correlation energies.
\(^3\) The standard deviation from the MP2/CBS values.
Table 2.4: The MC-MP2-F12(V) and MC-MP2-F12(VBX) correlation energies in $E_h$. The values in parentheses are statistical uncertainties.

<table>
<thead>
<tr>
<th>Molecule</th>
<th>MC-MP2-F12(V)</th>
<th>MC-MP2-F12(VBX)</th>
<th>CBS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>ADZ</td>
<td>ATZ</td>
<td>ADZ</td>
</tr>
<tr>
<td>H$_2$</td>
<td>-0.0336(0)</td>
<td>-0.0341(0)</td>
<td>-0.0342(0)</td>
</tr>
<tr>
<td>CH$_2$</td>
<td>-0.1508(1)</td>
<td>-0.1554(1)</td>
<td>-0.1535(1)</td>
</tr>
<tr>
<td>CH$_4$</td>
<td>-0.2134(1)</td>
<td>-0.2189(2)</td>
<td>-0.2165(1)</td>
</tr>
<tr>
<td>NH$_3$</td>
<td>-0.2601(1)</td>
<td>-0.2650(2)</td>
<td>-0.2603(2)</td>
</tr>
<tr>
<td>H$_2$O</td>
<td>-0.3004(1)</td>
<td>-0.3023(1)</td>
<td>-0.2954(2)</td>
</tr>
<tr>
<td>C$_2$H$_2$</td>
<td>-0.3263(1)</td>
<td>-0.3231(1)</td>
<td>-0.3129(2)</td>
</tr>
<tr>
<td>C$_2$H$_4$</td>
<td>-0.3337(1)</td>
<td>-0.3444(2)</td>
<td>-0.3389(3)</td>
</tr>
<tr>
<td>HF</td>
<td>-0.3622(2)</td>
<td>-0.3714(5)</td>
<td>-0.3673(4)</td>
</tr>
<tr>
<td>HNC</td>
<td>-0.3660(3)</td>
<td>-0.3739(3)</td>
<td>-0.3677(7)</td>
</tr>
<tr>
<td>HCN</td>
<td>-0.3774(2)</td>
<td>-0.3866(2)</td>
<td>-0.3789(4)</td>
</tr>
<tr>
<td>CO</td>
<td>-0.4012(3)</td>
<td>-0.4056(2)</td>
<td>-0.3970(5)</td>
</tr>
<tr>
<td>N$_2$</td>
<td>-0.4147(2)</td>
<td>-0.4218(2)</td>
<td>-0.4127(5)</td>
</tr>
<tr>
<td>CH$_2$O</td>
<td>-0.4439(3)</td>
<td>-0.4492(4)</td>
<td>-0.4405(5)</td>
</tr>
<tr>
<td>HNO</td>
<td>-0.4934(3)</td>
<td>-0.4987(3)</td>
<td>-0.4873(5)</td>
</tr>
<tr>
<td>H$_2$O$_2$</td>
<td>-0.5700(4)</td>
<td>-0.5727(5)</td>
<td>-0.5601(7)</td>
</tr>
<tr>
<td>F$_2$</td>
<td>-0.6253(3)</td>
<td>-0.6193(5)</td>
<td>-0.5978(7)</td>
</tr>
<tr>
<td>CO$_2$</td>
<td>-0.6828(6)</td>
<td>-0.6890(6)</td>
<td>-0.6730(12)</td>
</tr>
</tbody>
</table>

| Error$^4$ | 0.0066(3) | 0.0030(3) | 0.0069(5) | 0.0035(6) | 0.0000 |

$^1$ Geometries were taken from Ref.$^{91}$.  
$^2$ Only the F12 corrections were calculated by the MC-MP2-F12 method using either the nonvariational V formula or the variational VBX formula and the Slater-type geminal (geminal 1) with $\gamma = 1.1$ a.u. The number of MC steps was $1.44 \times 10^8$ and the number of the redundant walkers was 40.  
$^3$ The $X^{-3}$ extrapolation using the MP2/AQZ and MP2/A5Z correlation energies.  
$^4$ The standard deviation from the MP2/CBS values (the standard deviation of the statistical uncertainties).

formula, perhaps supporting its design.

2.4.3 Reaction energies

Energies of 12 gas-phase reactions listed in Table 2.6 were calculated by MP2, MP2-F12, and MC-MP2-F12 using geminal 1 (the Slater-type geminal) with $\gamma = 1.1$ a.u. The results are compiled in Tables 2.7 and 2.8 for the ADZ and ATZ basis sets respectively. They were computed as the difference in the total MP2 or MP2-F12 electronic energy between reactants and products, whose geometries were taken from Bak et al.$^{91}$ The HF and MP2 parts of the energy were obtained with the conventional deterministic algorithms in nwchem,$^{89}$ whereas the F12 part was computed either by MC-MP2-F12(V), MC-MP2-F12(VBX) or deterministic MP2-F12(VBX), the latter im-
Table 2.5: The average error (in \( mE_h \)) in the MC-MP2-F12 correlation energies using the geminals listed in Table 2.1 from the MP2/CBS values. The 17 molecules used were the same as those in Tables 2.3 and 2.4.

<table>
<thead>
<tr>
<th>Correlation factor</th>
<th>ADZ(^2)</th>
<th>ATZ(^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( f_{12}^{(1)} (\gamma = 1.1 \text{ a.u.}) )</td>
<td>( V ) 6.6 (0.3)</td>
<td>3.0 (0.3)</td>
</tr>
<tr>
<td></td>
<td>( \text{VBX} ) 6.9 (0.5)</td>
<td>3.5 (0.6)</td>
</tr>
<tr>
<td>( f_{12}^{(2)} (\gamma = 1.2 \text{ a.u.}) )</td>
<td>( V ) 20.2 (0.1)</td>
<td>17.0 (0.2)</td>
</tr>
<tr>
<td></td>
<td>( \text{VBX} ) 63.4 (0.2)</td>
<td>21.9 (0.3)</td>
</tr>
<tr>
<td>( f_{12}^{(3)} (\gamma = 1.1 \text{ a.u.}) )</td>
<td>( V ) 17.6 (0.3)</td>
<td>4.2 (0.3)</td>
</tr>
<tr>
<td></td>
<td>( \text{VBX} ) 8.7 (0.7)</td>
<td>4.0 (0.7)</td>
</tr>
<tr>
<td>( f_{12}^{(4)} (\gamma = 1.1 \text{ a.u.}) )</td>
<td>( V ) 17.6 (0.3)</td>
<td>5.2 (0.4)</td>
</tr>
<tr>
<td></td>
<td>( \text{VBX} ) 12.3 (0.7)</td>
<td>4.6 (0.7)</td>
</tr>
<tr>
<td>( f_{12}^{(5)} (\gamma = 1.1 \text{ a.u.}) )</td>
<td>( V ) 21.8 (0.2)</td>
<td>3.0 (0.3)</td>
</tr>
<tr>
<td></td>
<td>( \text{VBX} ) 12.1 (0.5)</td>
<td>3.2 (0.5)</td>
</tr>
<tr>
<td>( f_{12}^{(6)} (\gamma = 1.1 \text{ a.u.}) )</td>
<td>( V ) 12.8 (0.6)</td>
<td>2.1 (0.3)</td>
</tr>
<tr>
<td></td>
<td>( \text{VBX} ) 7.2 (0.6)</td>
<td>3.2 (0.6)</td>
</tr>
</tbody>
</table>

\(^1\) Only the F12 corrections were calculated by the MC-MP2-F12 method using either the nonvariational \( V \) formula or the variational \( VBX \) formula. The number of MC steps was 1.44 \( \times 10^8 \) and the number of the redundant walkers was 40.

\(^2\) The standard deviation from the MP2/CBS values (the standard deviation of the statistical uncertainties).

First, focusing on the ADZ results, we confirm that the reaction energies of MP2 suffer from excessively large errors from the CBS limits, which are, on average, 12.5 kJ mol\(^{-1}\). They are so large that even endo/exothermicity is incorrectly predicted (as compared with the CBS limits, if not with the experimental results) for reactions 2, 5, and 8. Once the F12 corrections from MC-MP2-F12 are added, be they based on the \( V \) or \( VBX \) formula, the reaction energies are, on average, within 2.6 kJ mol\(^{-1}\) of the CBS limits. The statistical uncertainties are no more than 1.7 kJ mol\(^{-1}\) after 1.44 \( \times 10^8 \) MC steps and are comparable to the intrinsic errors in the F12 method itself, but can be arbitrarily reduced by running longer MC integrations. Therefore, MC-MP2-F12 with either the \( V \) or \( VBX \) formula does work well in practice for relative energies, consistently achieving the chemical accuracy from the CBS limits.

We also find that MC-MP2-F12(\( VBX \)) gives the results that are systematically closer [than MC-MP2-F12(\( V \))] to deterministic MP2-F12(\( VBX \)) for all reactions barring one. Therefore, we argue that the MC-MP2-F12 method does not have a bias\(^88\) and is convergent at the correct limit, i.e., the result of the deterministic version of the corresponding method. The statistical uncertainties also seem correct in that in all cases the MC results are within 1\( \sigma \) (statistical uncertainty) from the deterministic values. The \( V \) results are often outside 3\( \sigma \) from the deterministic results, but
Figure 2.4: The percent recovery of the CBS limits of the MP2 correlation energies by various methods. The corresponding numerical data are in Tables 2.3 and 2.4.

this does not mean an underestimation of $\sigma$; it simply means that MC-MP2-F12(V) and MP2-F12(VBX) are two different methods with slightly different correct results. Note that deterministic MP2-F12(VBX) used here does not invoke either the GBC or EBC, unlike MC-MP2-F12(VBX), which assumes both. However, we numerically confirmed (not shown) that the differences in the F12 corrections caused by these approximations are much smaller than the typical statistical uncertainties and negligible, except when $\gamma$ is too small ($\gamma \leq 0.4$ a.u. in the case of the Slater-type geminal).

Furthermore, the errors from the CBS limits are distinctly smaller in MC-MP2-F12(VBX) than in MC-MP2-F12(V) with the ADZ basis set. This may well be because the intrinsic errors in the F12 method, which are always positive with the variational VBX formula, cancel between reactants and products in MC-MP2-F12(VBX). However, the statistical uncertainties are nearly twice as large in the VBX results as in the V results, making it difficult to draw a definitive conclusion. The greater statistical uncertainties in the results of the VBX formula are due to the larger number
Table 2.6: Gas-phase reactions considered in this work.

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Formula</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>CO + H₂ → CH₂O</td>
</tr>
<tr>
<td>2</td>
<td>N₂ + 3H₂ → 2NH₃</td>
</tr>
<tr>
<td>3</td>
<td>C₂H₂ + H₂ → C₂H₄</td>
</tr>
<tr>
<td>4</td>
<td>CO₂ + 4H₂ → CH₄ + 2H₂O</td>
</tr>
<tr>
<td>5</td>
<td>CH₂O + 2H₂ → CH₄ + H₂O</td>
</tr>
<tr>
<td>6</td>
<td>CO + 3H₂ → CH₄ + H₂O</td>
</tr>
<tr>
<td>7</td>
<td>HCN + 3H₂ → CH₄ + NH₃</td>
</tr>
<tr>
<td>8</td>
<td>HNO + 2H₂ → NH₃ + H₂O</td>
</tr>
<tr>
<td>9</td>
<td>C₂H₂ + 3H₂ → 2CH₄</td>
</tr>
<tr>
<td>10</td>
<td>CH₂ + H₂ → CH₄</td>
</tr>
<tr>
<td>11</td>
<td>F₂ + H₂ → 2HF</td>
</tr>
<tr>
<td>12</td>
<td>2CH₂ → C₂H₄</td>
</tr>
</tbody>
</table>

¹Geometries were taken from Ref. ⁹¹.

and higher dimension of its integrals, and are expected. Nonetheless, the difference in the overall accuracy between V and VBX is small, and given its much smaller computational cost (see Sec. 2.4.4) and somewhat smaller statistical uncertainty, the nonvariational V formula may be preferred in most applications.

Turning to the results of the ATZ basis set, we find the benefit of the F12 method to diminish rather rapidly with a basis-set extension especially in the MC-MP2-F12 implementation. The average error from the CBS limits is compressed significantly from 12.5 to 3.5 kJ mol⁻¹ going from MP2/ADZ to MP2/ATZ. In the meantime, the same quantity of MC-MP2-F12(V) decreases from 2.6 to 1.3 kJ mol⁻¹, while the average statistical uncertainty increases from 0.9 to 1.2 kJ mol⁻¹. With MC-MP-F12(VBX), the average error from the CBS limits increases from 1.7 kJ mol⁻¹ (ADZ) to 2.6 kJ mol⁻¹ (ATZ), probably owing to the increased statistical uncertainty from 1.7 to 2.3 kJ mol⁻¹. However, we do not consider this to be particularly troubling because it is with a small basis set that the F12 method is most needed. That MC-MP2-F12 is relatively more effective and efficient (in the sense of giving smaller statistical uncertainties) with ADZ than with ATZ should indeed be considered as a practical advantage.

Tables 2.9 and 2.10 summarize the geminal-dependence of the reaction energies calculated by MC-MP2-F12/ADZ with the V and VBX formulas respectively. Figure 2.5 plots the same using the VBX data and the values from the deterministic counterparts. The statistical uncertainties are comparable across different geminals, but the average errors from the CBS limits vary more greatly and are a better gauge of the geminals’ performance.

The Slater-type geminal (geminal 1) is confirmed to be the one with the smallest average errors,
Table 2.7: The MP2 reaction energies in kJ mol$^{-1}$ for the ADZ basis set and CBS. The values in parentheses are statistical uncertainties.

<table>
<thead>
<tr>
<th>Reaction</th>
<th>MP2</th>
<th>F12(VBX)$^3$</th>
<th>ADZ</th>
<th>MC-F12(V)$^4$</th>
<th>MC-F12(VBX)$^4$</th>
<th>CBS$^2$</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>−16.1</td>
<td>−24.6</td>
<td>−23.8 (1.0)</td>
<td>−24.2 (1.9)</td>
<td>−26.4</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>1.8</td>
<td>−13.3</td>
<td>−12.1 (0.8)</td>
<td>−13.5 (1.5)</td>
<td>−14.2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>14.2</td>
<td>16.2</td>
<td>13.6 (0.7)</td>
<td>15.4 (1.3)</td>
<td>17.2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>21.2</td>
<td>5.2</td>
<td>8.4 (1.7)</td>
<td>7.3 (3.3)</td>
<td>6.1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>0.5</td>
<td>−8.4</td>
<td>−6.8 (0.8)</td>
<td>−7.5 (1.6)</td>
<td>−8.4</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>−15.7</td>
<td>−33.1</td>
<td>−30.6 (0.7)</td>
<td>−31.7 (1.4)</td>
<td>−34.8</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>24.1</td>
<td>13.7</td>
<td>12.6 (0.6)</td>
<td>12.7 (1.1)</td>
<td>14.5</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>12.1</td>
<td>−0.1</td>
<td>0.5 (0.8)</td>
<td>0.3 (1.5)</td>
<td>−0.7</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>24.1</td>
<td>23.9</td>
<td>20.4 (0.6)</td>
<td>22.8 (1.1)</td>
<td>25.2</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>−65.8</td>
<td>−75.4</td>
<td>−76.2 (0.3)</td>
<td>−75.4 (0.5)</td>
<td>−76.0</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>28.3</td>
<td>17.4</td>
<td>16.4 (1.0)</td>
<td>16.1 (2.0)</td>
<td>16.7</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>−141.6</td>
<td>−158.4</td>
<td>−159.2 (0.6)</td>
<td>−158.3 (1.3)</td>
<td>−160.0</td>
<td></td>
</tr>
</tbody>
</table>

| Error$^3$ | 11.9 | 0.0  | 1.9 (0.9) | 1.0 (1.7) | 1.1 |
| Error$^6$ | 12.5 | 1.1  | 2.6 (0.9) | 1.7 (1.7) | 0.0 |

$^1$ See Table 2.6.
$^2$ See footnote c of Table 2.3.
$^3$ The deterministic MP2-F12 calculation using the variational VBX formula and the Slater-type geminal (geminal 1) with $\gamma = 1.1$ a.u. Unlike MC-MP2-F12, neither the GBC nor EBC was assumed.
$^4$ MC-MP2-F12 calculations. See footnote b of Table 2.3.
$^5$ The standard deviation from the deterministic MP2-F12(VBX)/ADZ values (the standard deviation of the statistical uncertainties).
$^6$ The standard deviation from the MP2/CBS values (the standard deviation of the statistical uncertainties).

either with the $V$ or $VBX$ formula, and is, therefore, judged to be the most suitable for the fixed-amplitude MP2-F12 method. However, insofar as the short-range $r_{12}$-dependence is similar to the Slater-type geminal (see Fig. 2.1), other geminals (barring geminal 2) work almost as effectively. For instance, geminal 3 (the rational geminal) has the average errors that are roughly twice those of geminal 1 (the Slater-type geminal). Geminal 6 (the hybrid geminal) is an average of geminals 1 and 3 and its errors also come in between those of its parent geminals. This geminal may be said to have the advantage that its F12 corrections are nearly independent of $\gamma$ (see Figs. 2.2 and 2.3). If an optimal value of $\gamma$ for a given geminal is unknown, one may use either the $VBX$ formula or geminal 6.
Table 2.8: The MP2 reaction energies in kJ mol\(^{-1}\) for the ATZ basis set and CBS. The values in parentheses are statistical uncertainties.

<table>
<thead>
<tr>
<th>Reaction</th>
<th>MP2</th>
<th>ATZ</th>
<th>MC-F12(V) (^4)</th>
<th>MC-F12(VBX) (^4)</th>
<th>CBS (^2)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-24.9</td>
<td>-24.9 (1.2)</td>
<td>-24.5 (2.3)</td>
<td>-26.4</td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>-12.2</td>
<td>-15.3 (1.0)</td>
<td>-18.2 (1.9)</td>
<td>-14.2</td>
<td></td>
</tr>
<tr>
<td>3</td>
<td>13.6</td>
<td>18.7 (1.4)</td>
<td>20.8 (2.6)</td>
<td>17.2</td>
<td></td>
</tr>
<tr>
<td>4</td>
<td>5.8</td>
<td>5.3 (1.9)</td>
<td>3.5 (3.6)</td>
<td>6.1</td>
<td></td>
</tr>
<tr>
<td>5</td>
<td>-7.9</td>
<td>-9.8 (1.2)</td>
<td>-11.9 (2.3)</td>
<td>-8.4</td>
<td></td>
</tr>
<tr>
<td>6</td>
<td>-32.8</td>
<td>-34.7 (0.9)</td>
<td>-36.5 (1.7)</td>
<td>-34.8</td>
<td></td>
</tr>
<tr>
<td>7</td>
<td>13.9</td>
<td>13.1 (0.9)</td>
<td>11.4 (1.7)</td>
<td>14.5</td>
<td></td>
</tr>
<tr>
<td>8</td>
<td>2.9</td>
<td>-0.9 (1.1)</td>
<td>-3.0 (2.0)</td>
<td>-0.7</td>
<td></td>
</tr>
<tr>
<td>9</td>
<td>21.3</td>
<td>23.3 (1.2)</td>
<td>23.5 (2.3)</td>
<td>25.2</td>
<td></td>
</tr>
<tr>
<td>10</td>
<td>-73.0</td>
<td>-77.2 (0.6)</td>
<td>-78.2 (1.1)</td>
<td>-76.0</td>
<td></td>
</tr>
<tr>
<td>11</td>
<td>22.2</td>
<td>18.9 (1.4)</td>
<td>18.3 (2.7)</td>
<td>16.7</td>
<td></td>
</tr>
<tr>
<td>12</td>
<td>-151.9</td>
<td>-159.1 (1.3)</td>
<td>-159.1 (2.5)</td>
<td>-160.0</td>
<td></td>
</tr>
<tr>
<td>Error(^3)</td>
<td>2.7</td>
<td>1.4 (1.2)</td>
<td>2.8 (2.3)</td>
<td>1.1</td>
<td></td>
</tr>
<tr>
<td>Error(^6)</td>
<td>3.5</td>
<td>1.3 (1.2)</td>
<td>2.6 (2.3)</td>
<td>0.0</td>
<td></td>
</tr>
</tbody>
</table>

\(^1\) See Table 2.6.
\(^2\) See footnote c of Table 2.3.
\(^3\) See footnote c of Table 2.7.
\(^4\) See footnote d of Table 2.7.
\(^5\) See footnote e of Table 2.7.
\(^6\) See footnote f of Table 2.7.

2.4.4 System-size scaling

Here, we determine the asymptotic functional dependence (scaling) of the cost of MC-MP2-F12 on the number of basis functions \((n)\), i.e., the system’s spatial size.

Figure 2.6 plots the cost scaling per MC step. To understand the observation, we briefly review the algorithm first. In each MC step, an electron-pair walker and \(m\) one-electron walkers need to be propagated by the Metropolis algorithm. The cost scaling of this step is \(O(mn)\) because a processor must evaluate \(O(n)\) weight functions at \(O(m)\) coordinates. Next, \(n\) AO amplitudes at \(m\) accepted walker coordinates are computed at an \(O(mn)\) cost. Then, these AO amplitudes are transformed into MO amplitudes at an \(O(mn^2)\) cost because at each of \(O(m)\) walker coordinates, \(O(n^2)\) multiplications of MO coefficients and AO amplitudes occur in this step. Thereupon, \(O_{pq}\) [Eq. (2.45)], \(V_{pq}\) [Eq. (2.46)], and \(O'_{pq}\) [Eq. (2.71)] arrays are constructed at an \(O(m^2n)\) cost, for each array is a sum over \(O(n)\) orbitals evaluated at two walker coordinates, whose number grows as \(O(m^2)\). Once these arrays are constructed, a processor accumulates the F12 correction. The cost
of this step is cheap and independent of \( n \) (but is heavily dependent on \( m \)).

Figure 2.6 shows that the wall time required for 128 MC steps of the MC-MP2-F12 calculations using the \( V \) formula falls accurately on the \( n^2 \) line in the large \( n \) limit. This is because the overall cost per MC step is dominated by the AO-to-MO transformation of orbital amplitudes, whose cost is an \( O(mn^2) \) quantity, when the number of redundant walkers (\( m \)) is as small as 40 in this case. For small \( n \), the cost per MC step appeared linear with \( n \),\(^5\) as the \( O(mn) \) AO amplitude calculation step is dominant with a large prefactor on the cost function caused by its many exponential evaluations. For larger \( n \), as explored in this work, the \( O(mn^2) \) MO amplitude calculation step supersedes the former, despite its small prefactor. When a large value of \( m \) is used, the cost per MC step reverts back to linear scaling with \( n \),\(^7\) since the \( O(m^2n) \) step of constructing \( O_{pq} \), \( V_{pq} \), and \( O'_{pq} \) becomes the hotspot. The same logic applies to the cost scaling of MC-MP2. Therefore, the cost scaling of MC-MP2-F12 varies with the balance between \( n \) and \( m \), but is observed to be \( O(n^2) \) per step in the present implementation and for a wide range of molecular sizes.

The cost function of the \( VBX \) formula may not have entered an asymptotic region by \( n = 472 \), but its functional form is almost certain to be \( O(n^2) \) because a \( VBX \) calculation includes all steps of the \( V \) calculation. For a small molecule, a \( VBX \) calculation is nearly an order of magnitude

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Slater (^1)</th>
<th>Gauss (^2)</th>
<th>Rational (^2)</th>
<th>Logarithm (^2)</th>
<th>Arctangent (^2)</th>
<th>Hybrid (^2)</th>
<th>CBS (^3)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>-23.8</td>
<td>-25.0</td>
<td>-23.6</td>
<td>-22.8</td>
<td>-23.3</td>
<td>-23.7</td>
<td>-26.4</td>
</tr>
<tr>
<td>2</td>
<td>-12.1</td>
<td>-23.4</td>
<td>-9.6</td>
<td>-6.5</td>
<td>-11.7</td>
<td>-9.4</td>
<td>-14.2</td>
</tr>
<tr>
<td>3</td>
<td>13.6</td>
<td>4.9</td>
<td>15.0</td>
<td>16.5</td>
<td>12.0</td>
<td>13.5</td>
<td>17.2</td>
</tr>
<tr>
<td>4</td>
<td>8.4</td>
<td>2.6</td>
<td>13.7</td>
<td>17.4</td>
<td>9.7</td>
<td>11.8</td>
<td>6.1</td>
</tr>
<tr>
<td>5</td>
<td>-6.8</td>
<td>-11.1</td>
<td>-3.5</td>
<td>-4.9</td>
<td>-5.2</td>
<td>-5.4</td>
<td>-8.4</td>
</tr>
<tr>
<td>6</td>
<td>-30.6</td>
<td>-36.1</td>
<td>-27.1</td>
<td>-27.7</td>
<td>-28.5</td>
<td>-29.1</td>
<td>-34.8</td>
</tr>
<tr>
<td>7</td>
<td>12.6</td>
<td>3.0</td>
<td>16.8</td>
<td>17.8</td>
<td>13.8</td>
<td>14.3</td>
<td>14.5</td>
</tr>
<tr>
<td>8</td>
<td>0.5</td>
<td>-13.8</td>
<td>4.6</td>
<td>5.6</td>
<td>1.2</td>
<td>1.6</td>
<td>-0.7</td>
</tr>
<tr>
<td>9</td>
<td>20.4</td>
<td>5.4</td>
<td>22.4</td>
<td>23.1</td>
<td>18.6</td>
<td>21.6</td>
<td>25.2</td>
</tr>
<tr>
<td>10</td>
<td>-76.2</td>
<td>-81.3</td>
<td>-74.3</td>
<td>-73.1</td>
<td>-75.2</td>
<td>-75.2</td>
<td>-76.0</td>
</tr>
<tr>
<td>11</td>
<td>16.4</td>
<td>-3.8</td>
<td>19.9</td>
<td>19.8</td>
<td>14.1</td>
<td>18.1</td>
<td>16.7</td>
</tr>
<tr>
<td>12</td>
<td>-159.2</td>
<td>-163.1</td>
<td>-156.1</td>
<td>-152.8</td>
<td>-156.9</td>
<td>-158.5</td>
<td>-160.0</td>
</tr>
</tbody>
</table>

Error \(^4\) | 2.6 | 10.9 | 4.5 | 5.7 | 3.8 | 3.4 | 0.0 |

---

\(^1\) See Table 2.6.
\(^2\) See footnote b of Table 2.3.
\(^3\) See footnote c of Table 2.3.

Table 2.9: The reaction energies in kJ mol\(^{-1}\) calculated by MC-MP2-F12(V)/ADZ using the geminals listed in Table 2.1.
Table 2.10: The reaction energies in kJ mol\(^{-1}\) calculated by MC-MP2-F12(VBX)/ADZ using the geminals listed in Table 2.1.

<table>
<thead>
<tr>
<th>Reaction</th>
<th>Slater(^2)</th>
<th>Gauss(^2)</th>
<th>Rational(^2)</th>
<th>Logarithm(^2)</th>
<th>Arctangent(^2)</th>
<th>Hybrid(^2)</th>
<th>CBS(^3)</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-24.2</td>
<td>-22.1</td>
<td>-26.7</td>
<td>-26.2</td>
<td>-25.6</td>
<td>-25.7</td>
<td>-26.4</td>
</tr>
<tr>
<td>2</td>
<td>-13.5</td>
<td>-19.8</td>
<td>-15.1</td>
<td>-12.7</td>
<td>-13.6</td>
<td>-11.7</td>
<td>-14.2</td>
</tr>
<tr>
<td>3</td>
<td>15.4</td>
<td>6.8</td>
<td>16.6</td>
<td>18.5</td>
<td>13.2</td>
<td>14.4</td>
<td>17.2</td>
</tr>
<tr>
<td>4</td>
<td>7.3</td>
<td>-0.1</td>
<td>14.5</td>
<td>15.3</td>
<td>7.4</td>
<td>10.7</td>
<td>6.1</td>
</tr>
<tr>
<td>5</td>
<td>-7.5</td>
<td>-10.9</td>
<td>-3.7</td>
<td>-9.6</td>
<td>-5.2</td>
<td>-6.0</td>
<td>-8.4</td>
</tr>
<tr>
<td>6</td>
<td>-31.7</td>
<td>-33.0</td>
<td>-30.3</td>
<td>-35.8</td>
<td>-30.8</td>
<td>-31.7</td>
<td>-34.8</td>
</tr>
<tr>
<td>7</td>
<td>12.7</td>
<td>8.1</td>
<td>16.0</td>
<td>14.7</td>
<td>13.9</td>
<td>13.1</td>
<td>14.5</td>
</tr>
<tr>
<td>8</td>
<td>0.3</td>
<td>-13.2</td>
<td>3.7</td>
<td>0.9</td>
<td>2.1</td>
<td>-0.2</td>
<td>-0.7</td>
</tr>
<tr>
<td>9</td>
<td>22.8</td>
<td>7.8</td>
<td>23.6</td>
<td>20.6</td>
<td>20.2</td>
<td>23.3</td>
<td>25.2</td>
</tr>
<tr>
<td>10</td>
<td>-75.4</td>
<td>-75.6</td>
<td>-75.7</td>
<td>-74.9</td>
<td>-76.1</td>
<td>-76.0</td>
<td>-76.0</td>
</tr>
<tr>
<td>11</td>
<td>16.1</td>
<td>-1.0</td>
<td>17.9</td>
<td>12.4</td>
<td>15.0</td>
<td>17.4</td>
<td>16.7</td>
</tr>
<tr>
<td>12</td>
<td>-158.3</td>
<td>-152.1</td>
<td>-158.4</td>
<td>-152.0</td>
<td>-159.3</td>
<td>-160.9</td>
<td>-160.0</td>
</tr>
<tr>
<td>Error(^4)</td>
<td>1.7</td>
<td>9.5</td>
<td>3.4</td>
<td>4.1</td>
<td>2.6</td>
<td>2.2</td>
<td>0.0</td>
</tr>
</tbody>
</table>

\(^1\) See footnote a of Table 2.9.  
\(^2\) See footnote b of Table 2.9.  
\(^3\) See footnote c of Table 2.9.  
\(^4\) See footnote d of Table 2.9.

more expensive than a \(V\) calculation. The gap becomes smaller and is only a factor of 2.3 for the largest molecule in the plot with \(n = 472\). Hence, in large molecules, the cost advantage of the \(V\) formula may be considered insignificant, rendering the \(V^\text{BX}\) formula preferred for its variational stability; recall that MC-MP2 and MC-MP2-F12 are designed for large molecules as they are not competitive with their deterministic counterparts for smaller ones.

The scaling of cost per MC step is not a useful measure of a stochastic method’s efficiency; the scaling of cost to achieve a given accuracy is. Figure 2.7 testifies that the relative statistical uncertainty (\(\sigma_{\text{rel}}\)) [as defined by Eq. (2.97)] after some MC steps (say, \(N_{\text{rel}}\)) is asymptotically proportional to \(n\):

\[
\sigma_{\text{rel}} = O(n).
\]  

(2.102)

As is well known, the statistical uncertainty in an MC integral tends to fall off accurately in proportion to \(N^{-1/2}\), where \(N\) is the number of MC steps. That this is also the case with the MC-MP2-F12 results has been confirmed with the data presented in Fig. 2.8 and elsewhere. Therefore,
Figure 2.5: The reaction energies obtained with the MP2, MP2-F12(VBX), and MC-MP2-F12(VBX) methods using the geminals listed in Table 2.1. The basis set is ADZ. The corresponding numerical data are found in Table 2.9 and 2.10. The error bars are the statistical uncertainties.

the relative statistical uncertainty ($\sigma_{\text{fin}}$) after $N_{\text{fin}}$ MC steps is accurately predicted to be

$$\sigma_{\text{fin}} \approx \frac{\sigma_{\text{rel}}}{\sqrt{N_{\text{fin}}/N_{\text{rel}}}} = \frac{O(n)}{\sqrt{N_{\text{fin}}}}. \quad (2.103)$$

For $\sigma_{\text{fin}}$ to become smaller than a given size-independent tolerance, $N_{\text{fin}}$ has to be an $O(n^2)$ quantity. The total cost is, therefore, $N_{\text{fin}}$ times the $O(n^2)$ cost per MC step, which is $O(n^4)$. We conclude that the MC-MP2-F12 cost to achieve a given accuracy (as measured by a relative statistical uncertainty) increases as $O(n^4)$. This may be compared with the $O(n^5)$ scaling of deterministic MP2 or MP2-F12 with fixed amplitudes.

2.4.5 Computer-size scaling

Figure 2.9 records the parallel scalability of MC-MP2-F12 on many CPUs. The speedup was measured as the rate of compression of the wall time spent in the last 25,000 MC steps of an MC-MP2-F12(VBX)/DZ calculation of the water molecule. The number of redundant walkers, $m$, was 20. The unit of speed was taken as that of an execution on 16 CPU cores or one XE node of Blue Waters at National Center for Supercomputing Applications of University of Illinois.

The MC-MP2-F12 algorithm is naturally parallel by design, in which each processor is tasked with its own Metropolis propagation of walkers and accumulation of the F12 correction with no mandatory interprocessor communications. All processors occasionally (every 128 MC steps) report their snapshot values of the F12 correction and statistical uncertainty to the master process.
Figure 2.6: The wall time (in seconds) spent in 128 MC steps of the MC-MP2-F12(VBX) (red dots) and MC-MP2-F12(V) (green dots) calculations as a function of the number of basis functions \(n\). The number of redundant walkers was 40. A set of 31 molecules ranging from water \((n = 10)\) to tetrahydrocannabinol \((n = 472)\) in the DZ basis set was used. Lines proportional to \(n\) and \(n^2\) are superimposed to guide the eyes.

by mpi_reduce, just to periodically protect the most current calculation results. As a result, it achieves scalability of 99.9% going from 16 CPU cores to 4,096 CPU cores or a 255.8-fold speedup for the tiniest problem of the water molecule in the DZ basis set, attesting to its strong scaling. For a larger calculation for tetrahydrocannabinol in the DZ basis set \((n = 472)\) shown in Fig. 2.8, a speedup by a factor of 7.84 (98% scalability) is observed from 512 to 4,096 CPU cores (using the VBX formula). Furthermore, calculations are indefinitely restartable in the sense that the F12 corrections and statistical uncertainties from two separate executions on different computers can be concatenated to produce the result for a longer continuous calculation.
Figure 2.7: The relative statistical uncertainty [Eq. (2.97)] of the MC-MP2-F12(VBX) (red dots) and MC-MP2-F12(V) (green dots) calculations as a function of the number of basis functions ($n$). The number of MC steps was $6.55 \times 10^6$, the number of redundant walkers was 40, and the block size $N_b$ was 7. The same set of molecules and basis set to generate Fig. 2.6 was used. Lines proportional to $n$ and $n^2$ are superimposed to guide the eyes.

### 2.5 Conclusions

Dense matrix multiplications, which have dominated electronic-structure algorithms, may be fundamentally nonscalable with both system and computer sizes. One may, therefore, need to redesign new, scalable algorithms when applications are attempted to large molecules, solids or even liquids, running on a modern supercomputer, which nowadays has up to hundreds of thousands or millions of processors. There are at least two such fundamentally scalable algorithms: local-basis (fragment or divide-and-conquer) algorithms\(^{101;102;103;104}\) and stochastic (Monte Carlo) algorithms.\(^{66;67;105}\)

In this work, we have fully developed a stochastic algorithm of MP2-F12 using both the variational VBX formula and nonvariational $V$ formula, the latter on the basis of the pilot implementa-
Figure 2.8: Convergence of the MP2 correlation energy of tetrahydrocannabinol (\(n = 472\)) in the DZ basis set with respect to the number of MC steps. The error bars correspond to \(\sigma_N\) of Eq. (2.96) with the block size of 7. The Slater-type geminal (geminal 1) with \(\gamma = 1.1\) a.u. was used. The number of redundant walkers was 40. The 4,096-way parallel \(V BX\) calculation with a total of \(7.87 \times 10^8\) MC steps took 7.17 hours of wall time.

On the basis of numerical experimentation, we have established the \(O(n^4)\) scaling of the MC-MP2-F12 method with the number of orbitals \(n\), which is one-rank lower than the usual \(O(n^5)\) scaling of MP2 or MP2-F12. The former does not need to use the RI approximation with an ABS to lower the dimensions of integrals or to evaluate the resulting long sum-of-products of lower-dimensional integrals, which is not scalable. It does not have to precompute or store any integrals and is free of any significant disk I/O.

Since all integrals (except one-electron integrals available from the HF calculations) are evaluated numerically by the MC method, MC-MP2-F12 can use nearly any mathematical form of geminal. Taking advantage of this, we have quantified the relative performance of the 6 geminals that satisfy the cusp conditions. We have confirmed the overall best performance of the Slater-type geminal. The Gauss-type geminal, which has a noticeably different long-range behavior than the
Figure 2.9: Speedup (relative to the 16-CPU run) of parallel MC-MP2-F12(VBX)/DZ calculations for the water molecule as a function of the number of CPU cores measured in terms of the wall time spent for the same number of samples. A speedup line for the perfect scalability is superimposed.

rest, is distinctly poorly performing to the extent that the reaction energies are hardly improved by its F12 correction.

Contrary to our concern that the nonvariational formula may be practically useless for relative energies because its F12 corrections are dependent on the shape of a geminal ($\gamma$, in particular) in a manner that is hard to predict, we found it to give reliable results at a computational cost that is 2–10 times smaller than the variational formula. Nevertheless, the relative energies from the variational formula are systematically more accurate because the errors caused by the suboptimal nature of the geminal shape cancel between the two total energies. For larger molecules, the cost differential between the two formulas becomes smaller, and it may thus be recommended to use the variational formula whenever possible. The statistical uncertainties in reaction energies are less than 2 kJ mol$^{-1}$ after $1.44 \times 10^8$ MC steps with the ADZ basis set, which are roughly half of the systematic errors of the F12 method itself (from the CBS limits) with the same basis set. The
former can be made arbitrarily smaller by running longer MC calculations.

The parallel scalability is excellent by virtue of the embarrassingly parallel nature, by design, of the algorithm. Each processor carries out its own MC integrations with no mandatory, frequent, or large interprocessor communications. A speedup by a factor of 255.8 has been achieved upon increasing the number of processors from 16 to 4,096, which is 99.9% of the perfect parallel efficiency, which may not be surprising when this method is viewed as a QMC variant. Our algorithm’s speed is likely limited only by the hardware size rather than by the software limitations.

MC-MP2-F12 and its sister methods are *not* intended to be a replacement of its deterministic counterparts and are slower and less precise than the latter for smaller problems. Rather, these two classes of implementations have completely different sets of merits and demerits, making them complementary in their applicabilities. MC-MP2-F12 is designed for grand-challenge applications on high-end massively parallel supercomputers, for which existing deterministic implementations may not even start owing to their more rigid resource requirements, higher cost scaling, and lower parallel scaling. Next, we will examine applicability and performance of MC-MP2-F12 for such large problems.
Chapter 3: Explicit correlation factors

This chapter is derived from joint work published with S. Hirata and S. Ten-no.\textsuperscript{106}

3.1 Introduction

The explicitly correlated extension\textsuperscript{32,51,78,79} of an \textit{ab initio} electron correlation method captures a much greater portion of the exact correlation energy with the aid of a correlation factor (a function of the interelectronic distance, $r_{12}$) than the same correlation method using an orbital basis set alone. Kutzelnigg\textsuperscript{18} introduced the so-called R12 ansatz of explicitly correlated method, using the simplest correlation factor that satisfies Kato’s cusp condition\textsuperscript{17}, namely, the one that is linear in $r_{12}$. Klopper and Kutzelnigg\textsuperscript{107} developed the first generally applicable R12 method with second-order many-body perturbation (MP2) theory using Gaussian-type orbital (GTO) basis sets. Insofar as the GTO basis set is large (so as to yield, e.g., 87\% of the exact MP2 limit for Ne), MP2-R12 is shown to recover an even greater portion of the exact correlation energy (99\% for Ne).

Ten-no\textsuperscript{20,50,99} proposed the use of the Slater-type geminal (STG), which has the form $(1 - e^{-\gamma r_{12}})/\gamma$ and thus satisfies the cusp condition\textsuperscript{17} at $r_{12} = 0$, but becomes flat as $r_{12} \to \infty$. This was shown to systematically and significantly outperform the linear factor\textsuperscript{20,50,99}. Tew and Klopper\textsuperscript{22} considered the Slater–linear factor, error function factor, error function–linear factor, and observed their equally good (if not better) performance. In analogy with the contracted GTO approximation of a Slater-type orbital, Valeev\textsuperscript{24} explored a correlation factor that was a linear combination of Gaussian-type geminals (GTG), and observed the near-optimal performance of a single STG. Indeed, a STG seems to resemble the cusped hole of a pair function (see, e.g., Ref.\textsuperscript{85}) of the Hylleraas functional so closely that its associated excitation amplitudes (geminal amplitudes) can be held fixed (the SP ansatz\textsuperscript{50}) at values dictated by the singlet and triplet cusp conditions\textsuperscript{17,48,80}. Today, a nonlinear correlation factor such as STG has completely replaced the linear factor in explicitly correlated methods, and those using the former are distinguished from the R12 methods as the F12 methods.

From these studies, it is evident that the shape of a correlation factor \textit{away from} $r_{12} = 0$ is equally important for the performance as the shape \textit{at} $r_{12} = 0$. Exact pointwise satisfaction of the cusp condition seems immaterial\textsuperscript{108}, judging from the fact that a GTG expansion of a STG works well\textsuperscript{22,24} despite GTG’s violation of the cusp condition. Klopper \textit{et al.}\textsuperscript{51} speculated the importance
of a nonzero second derivative of the correlation factor with respect to \( r_{12} \), which the linear factor lacks. However, the value of the second derivative (i.e., the second-order cusp condition\(^{109}\)) is system- and state-dependent and is not known \textit{a priori}, making it hard to judge the correctness of this speculation. Rassolov and Chipman\(^{110}\) derived an equality obeyed by the second and third derivatives, the significance of which is yet to be determined.

Monkhorst\(^{96}\) also analyzed the asymptotic forms of pair functions under various conditions and argued that two electrons should ‘decouple’ at large \( r_{12} \). This may explain the excellent performance of STG, which becomes constant at \( r_{12} \to \infty \), and the poor performance of the linear factor, which grows indefinitely. On the other hand, the correct long-range asymptote in a homogeneous electron gas was shown to be a screened Coulomb decay, supporting the Yukawa–Coulomb factor\(^{95}\), whose superior performance was numerically demonstrated by Grüneis \textit{et al.}\(^{97,111}\). In quantum Monte Carlo, which has extraordinary flexibility in the choice of correlation factors, the Jastrow factor\(^{94}\) is preferred, which differs from either STG or Yukawa–Coulomb.

In this study, we exploit the same extraordinary flexibility of the Monte Carlo MP2-F12 (MC-MP2-F12) method\(^{44,56}\) to numerically compare the performance of 17 correlation factors for small molecules. On this basis, we investigate the validity of the aforementioned claims concerning the shape of the correlation factor in determining the performance of the F12 method in the SP ansatz.

### 3.2 Computational details

Using the SP ansatz\(^{20,50}\) and generalized and extended Brillouin conditions\(^{21,49}\), the F12 correction\(^{18,77}\) to the MP2 energy consists of two parts:

\[
E_{\text{F12}}^{\text{V+BX}} = 2E_{\text{F12}}^{\text{V}} + E_{\text{F12}}^{\text{BX}}
\]

with

\[
E_{\text{F12}}^{\text{V}} = \frac{5}{8} \sum_{i,j} \text{occ} \ V_{ij}^{ij} - \frac{1}{8} \sum_{i,j} \text{occ} \ V_{ji}^{ij},
\]

\[
E_{\text{F12}}^{\text{BX}} = \frac{7}{32} \sum_{i,j} \text{occ} \ (BX)_{ij}^{ij} + \frac{1}{32} \sum_{i,j} \text{occ} \ (BX)_{ji}^{ij},
\]
and

\[ V_{mn}^{ij} = \langle ij | r_{12}^{-1} \hat{Q}_{12} f_{12} | mn \rangle, \quad (3.4) \]
\[ (BX)_{mn}^{ij} = \langle ij | f_{12} \hat{Q}_{12} [ \hat{F}_1 + \hat{F}_2, f_{12} ] | mn \rangle, \quad (3.5) \]

where \( f_{12} \) is the correlation factor, \( \hat{F}_n \) is the Fock operator for electron \( n \), and \( \hat{Q}_{12} \) is the strong-orthogonality projector:

\[ \hat{Q}_{12} = (1 - \hat{O}_1)(1 - \hat{O}_2) - \hat{V}_1 \hat{V}_2, \quad (3.6) \]

in the standard notation. The \( VBX \) expression [Eq. (3.1)] is variational with respect to the size and shape of the correlation factor. At its minimum, \( E_{F12}^B + E_{F12}^X = -E_{F12}^V \) and, therefore,

\[ E_{F12}^{VBX} = E_{F12}^V. \quad (3.7) \]

The overall size of the correlation factor (the geminal amplitudes) is held fixed in the SP ansatz. However, the shape of the correlation factor can usually be adjusted by varying some parameters, which are denoted by \( \gamma \) and \( \beta \) in this work. Equation (3.7) is expected to hold accurately at the minimum of \( E_{F12}^{VBX} \) with respect to \( \gamma \) and \( \beta \) if and only if the shape of the correlation factor is sufficiently close to that of the cusped hole of a pair function.

The total \( VBX \) energy can be divided into a sum over orbital-pair contributions as follows:

\[ E_{F12}^{VBX} = \sum_{i \leq j}^{\text{occ}} e_{ij} \quad (3.8) \]

with

\[ e_{ij} = (2 - \delta_{ij}) \left[ \frac{5}{4} V_{ij}^{ij} - \frac{1}{4} V_{ji}^{ij} + \frac{7}{32} (BX)^{ij}_{ij} + \frac{1}{32} (BX)^{ij}_{ji} \right]. \quad (3.9) \]

In this work, the high-dimensional integrals arising from Eqs. (3.2) and (3.3)\(^{56} \) were evaluated by the Monte Carlo (MC) method\(^ {44,56} \). It has the unique ability to handle virtually any integrand and thus any correlation factor, which is essential for our purpose. More specifically, the \( VBX \) formalism can treat any correlation factor that is analytic (having at least first and second derivatives with respect to \( r_{12} \)), whereas the \( V \) formalism can use any factor including numerically defined ones. This contingency on existing derivatives derives from the kinetic energy operator present in the B term of the \( VBX \) formalism.
We considered 17 analytic correlation factors in Table 3.1 including 6 that had been studied in Ref.\textsuperscript{56} for H\textsubscript{2}O and CH\textsubscript{4} using the cc-pVDZ basis set. The MC-MP2-F12 calculations used the implementation reported in Ref.\textsuperscript{56} with 8.65 × 10\textsuperscript{7} MC steps (3.74 × 10\textsuperscript{6} steps for Fig. 3.4), 40 redundant walkers for both one-electron and two-electron walkers\textsuperscript{113}, and 6 blocking transformations\textsuperscript{87} for statistical uncertainties. In all cases, the statistical uncertainties were small enough to not alter the conclusions drawn below (the statistical uncertainty with 99% confidence for the CBS recovery ratios was 0.1–0.4% with the maximum being 0.43% for correlation factor 3). The MP2 energies with the cc-pVDZ basis set and in the complete-basis-set (CBS) limits (see Ref.\textsuperscript{56} for more details) were obtained with \texttt{nwchem}\textsuperscript{89}. The core electrons were excluded in correlation calculations.

### 3.3 Results

Table 3.1 compiles the CBS recovery ratios (averaged over H\textsubscript{2}O and CH\textsubscript{4}) for all 17 correlation factors at values of $\gamma$ and $\beta$ that are near optimal (which are also listed in the table for H\textsubscript{2}O). For correlation factor 17, the results for 3 different values of $\beta$ are given. The first, second, and third derivatives of the correlation factor with respect to $r_{12}$ at $r_{12} = 0$ are also shown. The value of unity for the first derivative means that the correlation factor satisfies Kato’s cusp condition; only GTG (besides null correlation factor listed as $f_{12}^{(0)}$) does not satisfy the condition. Correlation factor 17 (STG–cubic exponential) reduces to a single STG as $\beta \to 0$. With decreasing $\beta$, the performance (CBS recovery ratio) of this correlation factor increases toward that of STG. This means that the cubic exponential part of this correlation factor acts to only spoil the high performance of STG.

Figure 3.1 plots the CBS recovery ratios for H\textsubscript{2}O and CH\textsubscript{4} separately for the correlation factors listed in Table 3.1. The ratios for CH\textsubscript{4} are higher than those for H\textsubscript{2}O by a few percents, the cause of which is unknown. From this figure and Table 3.1, we divide the correlation factors into two groups: the highly performing ones that capture 90% or above of the correlation energies in the CBS limit, and the poorly performing ones that fail to do so. The poorly performing factors are 2 (cusped Gaussian), 7 (linear), 9 (GTG), and 15 (error function–linear), while the rest are in the highly performing group. For this small dataset, the 3 best performing factors are 8 and 10 (tie by higher rational and Yukawa–Coulomb), 13 (higher Jastrow), 1 and 6 (tie by STG and hybrid) in this order, although the differences are not so meaningful in light of the errors of a few tenths of 1% from statistical uncertainties of MC integrations.

Figure 3.2 plots the correlation factors in these two groups as a function of $r_{12}$. Note that the values of $\gamma$ used in this plot are near optimal for H\textsubscript{2}O (given in Table 3.1). Common fea-
### Table 3.1: Correlation factors.

<table>
<thead>
<tr>
<th>Correlation factor</th>
<th>Name</th>
<th>$\gamma$ / a.u.$^1$</th>
<th>$\beta$ / a.u.$^1$</th>
<th>%CBS$^2$</th>
<th>$\partial f_{12}/\partial r_{12}$</th>
<th>$\partial^2 f_{12}/\partial r_{12}^2$ / a.u.</th>
<th>$\partial^3 f_{12}/\partial r_{12}^3$ / a.u.</th>
</tr>
</thead>
<tbody>
<tr>
<td>$f_{12}^{(0)} = 0$</td>
<td>None</td>
<td>...</td>
<td>...</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$f_{12}^{(1)} = (1 - e^{-\gamma r_{12}})/\gamma$</td>
<td>Slater (STG)</td>
<td>1.2</td>
<td>...</td>
<td>70.4</td>
<td>0</td>
<td>$-\gamma = 1.2$</td>
<td>$\gamma^2 = 1.4$</td>
</tr>
<tr>
<td>$f_{12}^{(2)} = (1 - e^{-\gamma r_{12}})/(\gamma r_{12})$</td>
<td>Cusped Gaussian</td>
<td>1.2</td>
<td>...</td>
<td>82.4</td>
<td>1</td>
<td>0</td>
<td>$-3\gamma = -3.6$</td>
</tr>
<tr>
<td>$f_{12}^{(3)} = \gamma r_{12}/(\gamma + r_{12})$</td>
<td>Rational</td>
<td>1.2</td>
<td>...</td>
<td>95.9</td>
<td>1</td>
<td>$-2/\gamma = -1.7$</td>
<td>$6/\gamma^2 = 4.2$</td>
</tr>
<tr>
<td>$f_{12}^{(4)} = \ln(1 + \gamma r_{12})/\gamma$</td>
<td>Logarith</td>
<td>2.0</td>
<td>...</td>
<td>95.3</td>
<td>1</td>
<td>$-\gamma = -2.0$</td>
<td>$2\gamma^2 = 8.0$</td>
</tr>
<tr>
<td>$f_{12}^{(5)} = \arctan(\gamma r_{12})/\gamma$</td>
<td>Arctangent</td>
<td>1.6</td>
<td>...</td>
<td>96.0</td>
<td>1</td>
<td>0</td>
<td>$-2\gamma^2 = -5.1$</td>
</tr>
<tr>
<td>$f_{12}^{(6)} = f_{12}^{(1)} + f_{12}^{(3)}/2$</td>
<td>Hybrid</td>
<td>1.2</td>
<td>...</td>
<td>96.1</td>
<td>1</td>
<td>$-1/\gamma - \gamma/2 = -1.4$</td>
<td>$3/\gamma^2 + \gamma^2/2 = 2.8$</td>
</tr>
<tr>
<td>$f_{12}^{(7)} = r_{12}$</td>
<td>Linear</td>
<td>...</td>
<td>...</td>
<td>21.3</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$f_{12}^{(8)} = \gamma r_{12}/(2\gamma + 2r_{12})$</td>
<td>Higher rational</td>
<td>1.6</td>
<td>3.0</td>
<td>96.4</td>
<td>1</td>
<td>$-2/\beta - 1/\gamma = -1.3$</td>
<td>$9/\beta^2 + 3/\gamma^2 = 2.2$</td>
</tr>
<tr>
<td>$f_{12}^{(9)} = 1 - e^{-\gamma r_{12}}$</td>
<td>Gaussian (GTG)</td>
<td>0.5</td>
<td>...</td>
<td>80.9</td>
<td>0</td>
<td>$2\gamma = 1.0$</td>
<td>0</td>
</tr>
<tr>
<td>$f_{12}^{(10)} = (2e^{-\gamma r_{12}} - 2)/(\gamma^2 r_{12})$</td>
<td>Yukawa–Coulomb</td>
<td>2.0</td>
<td>...</td>
<td>96.4</td>
<td>1</td>
<td>$-2\gamma/3 = -1.3$</td>
<td>$\gamma^2 / 2 = 2.0$</td>
</tr>
<tr>
<td>$f_{12}^{(11)} = r_{12} e^{-\gamma r_{12}}$</td>
<td>Slater–linear</td>
<td>0.5</td>
<td>...</td>
<td>94.2</td>
<td>1</td>
<td>$-2\gamma = -1.0$</td>
<td>$3\gamma^2 = 0.8$</td>
</tr>
<tr>
<td>$f_{12}^{(12)} = e^{\gamma r_{12}^{(1)}/(1+\gamma r_{12})}$</td>
<td>Jastrow</td>
<td>1.2</td>
<td>...</td>
<td>96.0</td>
<td>1</td>
<td>$1 - 2\gamma = -1.4$</td>
<td>$1 - 6\gamma + 6\gamma^2 = 2.4$</td>
</tr>
<tr>
<td>$f_{12}^{(13)} = e^{U_{12}/(1+\gamma r_{12})}$</td>
<td>Higher Jastrow</td>
<td>0.8</td>
<td>0.75</td>
<td>96.2</td>
<td>1</td>
<td>$1 - \beta - 2\gamma = -1.4$</td>
<td>$1 - 6\gamma + 6\gamma^2 + \beta^2 + \beta(6\gamma - 3) = 2.0$</td>
</tr>
<tr>
<td>$U_{12} = (1 - e^{-\beta r_{12}})\beta$</td>
<td>Error function</td>
<td>1.2</td>
<td>...</td>
<td>92.6</td>
<td>1</td>
<td>0</td>
<td>$-2\gamma^2 = -2.9$</td>
</tr>
<tr>
<td>$f_{12}^{(14)} = -\sqrt{\pi} \text{erfc}(\gamma r_{12})/(2\gamma)$</td>
<td>Error function–linear</td>
<td>0.4</td>
<td>...</td>
<td>88.2</td>
<td>1</td>
<td>$-4\gamma/\pi = -0.5$</td>
<td>0</td>
</tr>
<tr>
<td>$f_{12}^{(15)} = r_{12} \text{erfc}(\gamma r_{12})$</td>
<td>Hyperbolic tangent</td>
<td>1.2</td>
<td>...</td>
<td>94.1</td>
<td>1</td>
<td>0</td>
<td>$-2\gamma^2 = -2.9$</td>
</tr>
<tr>
<td>$f_{12}^{(16)} = (1 - e^{-\gamma r_{12}})/\gamma + e^{-\beta r_{12}}$</td>
<td>Slater–cubic exponential</td>
<td>1.2</td>
<td>0.003</td>
<td>94.9</td>
<td>1</td>
<td>$-\gamma = -1.2$</td>
<td>$\gamma^2 - 6\beta = 1.42$</td>
</tr>
<tr>
<td>$f_{12}^{(17)} = (1 - e^{-\gamma r_{12}})/(\gamma + e^{-\beta r_{12}})$</td>
<td>Slater–cubic exponential</td>
<td>1.2</td>
<td>0.006</td>
<td>93.5</td>
<td>1</td>
<td>$-\gamma = -1.2$</td>
<td>$\gamma^2 - 6\beta = 1.40$</td>
</tr>
<tr>
<td>$f_{12}^{(18)} = (1 - e^{-\gamma r_{12}})/(\gamma + e^{-\beta r_{12}})$</td>
<td>Slater–cubic exponential</td>
<td>1.2</td>
<td>0.012</td>
<td>91.3</td>
<td>1</td>
<td>$-\gamma = -1.2$</td>
<td>$\gamma^2 - 6\beta = 1.37$</td>
</tr>
</tbody>
</table>

1 Near-optimal values of $\gamma$ and $\beta$ in the case of H$_2$O using the cc-pVDZ basis set.

2 The recovery ratio of the MP2 correlation energy in the complete-basis-set (CBS) limit averaged over H$_2$O and CH$_4$ using the cc-pVDZ basis set. Statistical uncertainties alone cause an error up to a few tenths of 1%.
Figure 3.1: The CBS recovery ratios of the MC-MP2-F12/cc-pVDZ calculations for H$_2$O and CH$_4$ using the correlation factors listed in Table 3.1 at their respective near-optimal $\gamma$ and $\beta$ values. Correlation factor 17 used $\beta = 0.003$ a.u.

features observed in the shapes of the highly performing correlation factors will likely reflect the near-universal shape of the cusped hole in a pair function, which cannot be described well by an orbital expansion. It is striking to observe the convergence of the functional forms of many best-performing correlation factors in the range of $0 < r_{12} < 1.5$ a.u. For instance, with increasing $\beta$, correlation factor 17 (STG–cubic exponential) more rapidly dives with $r_{12}$ and deviates more greatly from STG. As a result, this factor with $\beta = 0.012$ a.u. (the lowest-lying curve) does not differ much in the functional form from correlation factor 15 (error function–linear), which is among the poorly performing group. Correspondingly, its CBS recovery ratio is also the lowest among the highly performing group. As a side note, some of the correlation factors may have a fundamentally infinite variance due to their derivatives having higher order singularities in them beyond $r_{12}$, which are thus not canceled by the pair weight function. Because of this, some correlation factors may actually converge slower than $N^{-1/2}$ but this was not analyzed deeply in this study.
Figure 3.2: Plots of (a) poorly performing correlation factors and (b) highly performing ones at the values of $\gamma$ and $\beta$ given in Table 3.1. The blue-shaded area likely reflects the universal shape of the cusped holes of pair functions. Some curves are displaced vertically to cross the origin.
Figure 3.3 draws the MP2-F12 energies in the variational $VBX$ and nonvariational $V$ formalisms as a function of $\gamma$, complementing a similar figure for correlation factors 1 through 6 in Ref. 56. Correlation factor 7 (linear) is excluded because it has no $\gamma$ dependence, and the plot for correlation factor 13 (higher Jastrow) is also omitted as it is similar to the one for correlation factor 12 (Jastrow). We observe that with the poorly performing correlation factors 9 and 15, the $VBX$ and $V$ curves do not even cross in the chosen domain of $\gamma$. With the best performing correlation factors 10 (Yukawa–Coulomb) and 12 (Jastrow), in contrast, these two curves cross near the minima of the $VBX$ curves, satisfying Eq. (3.7). This, in turn, means that only when correlation factors have the right shape does the SP ansatz work exceedingly well. The appropriateness of the shape of a correlation factor can, therefore, be detected by the way these two curves intersect (in addition to the CBS recovery ratio). The $\beta$-dependence in the plot of correlation factor 17 (STG–cubic exponential) is a case in point. With decreasing $\beta$, this correlation factor approaches STG, improving its CBS recovery ratio. Correspondingly, the intersection of the $VBX$ and $V$ curves also shifts toward the minimum of the $VBX$ curve, attesting to the gradual improvement of the functional form.

### 3.4 Discussion

#### 3.4.1 Cusp ($r_{12} = 0$)

Prendergast et al. 108 argued that satisfying the cusp condition pointwise (at $r_{12} = 0$) alters the correlation energy only in the $\mu E_h$-accuracy range, and is rather insignificant, the assertion also supported (qualitatively) by Valeev’s work 24. Our results are not inconsistent with this conclusion in the sense that there are several correlation factors (such as 2, 7, and 15) that satisfy the cusp condition but perform extremely poorly in the SP ansatz. There is only one correlation factor (9 or GTG) in our list that does not satisfy the cusp condition. It also performs poorly, but this is likely traced to the wrong short-range shape (see below) rather than to the lack of cusp. Therefore, our results do not contain discerning (let alone quantitative) information about this conclusion, but they do not contradict it. Taken together with the other observations (see below) about the correct shape of the correlation factor, our results seem supportive of the notion that the pointwise (i.e., at $r_{12} = 0$) satisfaction of the cusp condition is not important for energetics, which may be self-evident because of the infinitesimally small volume element of the cusp. However, the cusp condition remains crucial as it dictates the correct short-range shape of a correlation factor near (if not at) the cusp (see below).

Figure 3.4 shows the breakdown of the F12 correction (using correlation factor 1 or STG)
Figure 3.3: The MC-MP2-F12/cc-pVDZ correlation energy as a function of $\gamma$ (and $\beta$) in the variational $VBX$ and nonvariational $V$ formalisms of $H_2O$. See Ref.\textsuperscript{56} for the corresponding plots for correlation factors 1 through 6.
into its orbital-pair contributions [Eqs. (3.8) and (3.9)] as a function of $\gamma$. It can be seen that all valence-orbital-pair F12 corrections have a minimum at a similar optimal value of $\gamma$ around 1.2 a.u.

### 3.4.2 Long-range shape ($r_{12} > 1.5$ a.u.)

Figure 3.2 rather decisively shows that correlation factors that are equally highly performing can have qualitatively different long-range behavior at $r_{12} > 1.5$ a.u. Some (e.g., correlation factor 4) continue to increase with $r_{12}$, while others become constant (e.g., correlation factor 1) or turn to decrease (e.g., correlation factor 11), with no drastic effect on the CBS recovery ratios. We, therefore, argue that the long-range behavior of a correlation factor is not too important for the performance of the correlation factor, insofar as it does not distort its behavior in the short range ($r_{12} < 1.5$ a.u.). This is consistent with Monkhorst’s argument\(^{96}\) that two electrons should decouple its correlation in the long range; a good correlation factor should be constant or only moderately dependent on $r_{12}$ in the long range, so as not to interfere with the ability of an orbital expansion to accurately describe pair functions there. The linear factor (7) is the worst performer in the SP ansatz partly because of its inappropriately strong coupling at large $r_{12}$.

The Yukawa–Coulomb factor (10), which is said to have the physically correct $-2/(\gamma^2 r_{12})$ decay in the homogeneous electron gas, is found to be the best performer, which may be taken to support the claim of Grünies et al.\(^{97,111}\) for molecules as well as for homogeneous electron gases. However, the higher rational (8), which is tied for the highest CBS recovery ratio, has the $\beta^2/(2 r_{12})$ long-range behavior, but with the opposite sign. Furthermore, the differences in the CBS recovery ratios between these two and STG (correlation factor 1, which becomes constant at $r_{12} \rightarrow \infty$) are small and not statistically significant. Therefore, we are inclined to stand by the conclusion about the relative insignificance of the long-range behavior at $r_{12} > 1.5$ a.u.

The long-range asymptote of a pair function is expected to depend on the highest occupied orbital energy and thus on the system\(^{96}\). The fact that the F12 method with the same correlation factor (with the same $\gamma$ in the SP ansatz) works well for a variety of molecules\(^{20,50,99}\) also suggests that the long-range part of pair functions is described well by an orbital expansion; a correlation factor should be near constant there not to spoil this expansion.

### 3.4.3 Short-range shape ($0 < r_{12} < 1.5$ a.u.)

The most striking result of this work is the convergence of all high-performance correlation factors in the short range ($0 < r_{12} < 1.5$ a.u.) as seen in Fig. 3.2(b). The dense manifold of curves (indicated by a blue shade) in Fig. 3.2(b) in this $r_{12}$ range should represent the common feature of
the cusped hole in a pair function with the radius of approximately 1.5 a.u., which is not dissimilar to the ones observed in grid-based MP2 calculations. Pair functions outside this range are more strongly dependent on system and state, but are reasonably accurately describable by an orbital expansion. A few curves that are slight outliers of the manifold, such as correlation factors 14 (error function) and 16 (hyperbolic tangent), are also the ones with slightly lower CBS recovery ratios (92.6% and 94.1%, respectively). Correlation factor 17 with the largest value of $\beta = 0.012$ a.u. seems to decay too early (at $r_{12} < 1.5$ a.u.), which may explain the penalty in its performance (the CBS recovery ratio of 91.3%). The poorly performing correlation factors in Fig. 3.2(a) have vastly different shapes in this $r_{12}$ range.

We, therefore, conclude that the correct shape of the correlation factor in $0 < r_{12} < 1.5$ a.u. is crucial for the performance of the F12 method in the SP ansatz.
3.4.4 Second-order cusp and Rassolov–Chimpan conditions

The correct shape of the correlation factor is cusped at \( r_{12} = 0 \) and concave with a radius of approximately 1.5 a.u. This is consistent with the speculation of Klopper et al.\(^\text{51}\) about the significance of a nonzero second derivative of the correlation factor with respect to \( r_{12} \). Table 3.1 indicates that the highly performing correlation factors tend to have the second derivatives of around \(-1.2\) to \(-1.4\) a.u. It is possible for a correlation factor to have the correct shape in the short range with zero second derivative by having a compensating third derivative (such as in correlation factors 5 and 16). However, all poorly performing factors have zero or positive second derivative, making their shapes too linear or convex in the short range.

The above observation notwithstanding, the second derivative does not obey a simple, universal condition independent of system or state\(^\text{109,110}\). Rassolov and Chipman\(^\text{110}\), however, found the following condition to be met by the second and third derivatives, if we ignore the \( r_{12} \) dependence in the Slater determinant:

\[
\frac{\partial^3 f_{12}}{\partial r_{12}^3} \bigg|_{r_{12}=0} = \frac{\partial^2 f_{12}}{\partial r_{12}^2} \bigg|_{r_{12}=0} - \frac{1}{4}.
\]

None of the correlation factors we studied satisfies this condition; in most cases, the left- and right-hand sides of the condition even have the opposite signs. This may mean that the third-order Taylor expansion of the correlation factor is too approximate and does not discern the short-range shape up to \( r_{12} = 1.5 \) a.u., not to mention whether it has the near-constant behavior in the long range. But given the complicated dependence of the condition on our correlation factor and Slater determinant and the CBS recovery derived for the correlation factors in this study, this condition can be safely ignored without meaningfully affecting quantitative accuracy to the CBS limit.

3.4.5 Orbital-pair dependence

Figure 3.4 testifies that the \( \gamma \) values at which valence-orbital-pair F12 corrections [\( \epsilon_{ij} \) of Eq. (3.9)] in \( \text{H}_2\text{O} \) are at minimum are essentially unchanged from one pair to another. A similar result was obtained for \( \text{CH}_4 \) (not shown). This suggests that cusped holes in pair functions have the universal shape seen in Fig. 3.2(b) across valence orbital pairs (although core orbital pairs may have tighter cusps\(^\text{20}\)). We argue that this relative invariance in the size and shape of the cusped holes is what makes the SP ansatz with a nonlinear correlation factor a great success. It may be recalled that the second-order cusp condition is system- and orbital-pair-dependent\(^\text{109}\), whereas the appropriate form of the correlation factor seems the same for all valence orbital pairs.
Chapter 4: MC-GF2-F12 method

This chapter is derived from joint work with S. Hirata, A. E. Doran and S. Ten-no that is in preparation.

4.1 Introduction

The previous chapters have focused specifically on the corrections to the basis-set convergence of MBPT and their stochastic implementation, as well as an exploration of many different correlation factors. We focus now on another extremely important class of methods, the Green’s functions methods for electron-binding energies. One-particle many-body Green’s function (MBGF) theory\textsuperscript{12,13,114,115,116,117,118,119} provides a converging series of perturbation approximations to electron-binding energies, which are key to many processes in chemistry, biology, condensed matter physics, and materials science. Green’s function also serves as the propagator in quantum field theory, underlying nearly all quantum many-body physics. Furthermore, it is the mathematical basis of quantum transport, scattering, and embedding theories.

General-order algorithms of the Feynman–Dyson perturbation series of MBGF and its convergence rate as well as the accuracy of the diagonal and frequency-independent approximations to the self-energy have been recently implemented.\textsuperscript{120} The study showed that the convergence is surprisingly slow, making higher-order corrections much more important than in many-body perturbation theory (MBPT) for the ground-state correlation energy. This slow convergence is worrisome in view of the high-rank polynomial size dependence of the costs of the higher-order methods.

The problem size is the product of the system’s spatial extent and basis-set size. The problem is compounded by the notoriously slow basis-set convergence\textsuperscript{51,54,78,79,121} of quantities obtained from any \textit{ab initio} electron-correlation theory. The latter convergence problem was, however, recently resolved by an explicitly correlated (F12) extension of the second-order MBGF in the diagonal, frequency-independent approximation to the self-energy (GF2-F12) implemented for massively parallel execution.\textsuperscript{29} Their formalism is derived directly from the MP2-F12 pair energies, and is therefore only valid for the ionization potentials (IPs).\textsuperscript{29}

In this article, we report a stochastic algorithm of GF2-F12 that reduces its scaling with system size while maintaining the parallel efficiency. It achieves these by combining the Monte Carlo second-order Green’s function (MC-GF2)\textsuperscript{58} and Monte Carlo explicitly correlated second-
order many-body perturbation (MC-MP2-F12) methods as well as the deterministic GF2-F12 method of Ohnishi and Ten-no. We show that the resulting method, Monte Carlo explicitly correlated second-order many-body Green’s function (MC-GF2-F12) method, yields electron-binding energies near their complete-basis-set (CBS) limits at a one-rank lower size dependence in a naturally parallel algorithm.

4.2 Theory

Following the derivation and notation of Ohnishi and Ten-no, we write the second-order perturbation-corrected energy of the $x$th occupied orbital, $\epsilon_{x}^{GF2}$, as a sum of three parts,

$$
\epsilon_{x}^{GF2} = \epsilon_{x}^{HF} + \Sigma_{x}^{OR} + \Sigma_{x}^{CD},
$$

(4.1)

where $\epsilon_{x}^{HF}$ is the canonical Hartree–Fock (HF) orbital energy and $\Sigma_{x}^{OR}$ and $\Sigma_{x}^{CD}$ are the so-called orbital/pair-relaxation and correlation-difference contributions, respectively, to the second-order self-energy in the diagonal, frequency-independent approximation. They are given as

$$
\Sigma_{x}^{OR} = \sum_{i,j}^{occ.} \sum_{a}^{vir.} \frac{\langle ij|xa\rangle(2\langle xalij \rangle - \langle axlij \rangle)}{\epsilon_{x}^{HF} + \epsilon_{a}^{HF} - \epsilon_{i}^{HF} - \epsilon_{j}^{HF}},
$$

(4.2)

$$
\Sigma_{x}^{CD} = \sum_{j}^{occ.} \sum_{a,b}^{vir.} \frac{\langle xj|ab\rangle(2\langle ab|xj \rangle - \langle ba|xj \rangle)}{\epsilon_{x}^{HF} + \epsilon_{i}^{HF} - \epsilon_{a}^{HF} - \epsilon_{b}^{HF}},
$$

(4.3)

where $i, j, k, l, m, n$ and $x$ denote an occupied spatial orbital in a closed-shell molecule, while $a$ and $b$ a virtual spatial orbital.

As shown in the previous study, $\Sigma_{x}^{OR}$ converges towards the CBS limit quickly, whereas the basis-set convergence of $\Sigma_{x}^{CD}$ is rather slow owing to its double virtual summation. Therefore, only $\Sigma_{x}^{CD}$ needs to be corrected for the basis-set incompleteness. Rewriting $\Sigma_{x}^{CD}$ as

$$
\Sigma_{x}^{CD} = \sum_{j}^{occ.} e_{xj},
$$

(4.4)

with

$$
e_{xj} = \sum_{a,b}^{vir.} \frac{\langle xj|ab\rangle(2\langle ab|xj \rangle - \langle ba|xj \rangle)}{\epsilon_{x}^{HF} + \epsilon_{i}^{HF} - \epsilon_{a}^{HF} - \epsilon_{b}^{HF}},
$$

(4.5)
we notice that second-order many-body perturbation energy and its F12-corrected value, \( E^{\text{MP2}} \) and \( E^{\text{MP2-F12}} \), are also written as

\[
E^{\text{MP2}} = \sum_{i,j} e_{ij}, \tag{4.6}
\]

\[
E^{\text{MP2-F12}} = \sum_{i,j} (e_{ij} + d_{ij}), \tag{4.7}
\]

where \( d_{ij} \) is the F12 correction to the second-order pair energy. Hence, the F12-corrected value of the GF2 energy, \( \epsilon^{\text{GF2-F12}} \), can be obtained as

\[
\epsilon^{\text{GF2-F12}} = \epsilon^{\text{HF}} + \sum \text{OR} + \sum \text{CD} - F_{12}.
\]

The F12 correction to GF2, \( d_{xj} \), can be derived directly from \( d_{ij} \) in MP2-F12 and is

\[
\Sigma^{\text{F12}} = 2\Sigma^V + \Sigma^B + \Sigma^X, \tag{4.10}
\]

\[
\Sigma^V = \sum_{i,m,n} V_{mn}^{xj} (2t_{mxj} - t_{jx}^m), \tag{4.11}
\]

\[
\Sigma^B = \sum_{i,k,l,m,n} B_{kl}^{ij} (2t_{mxj} - t_{jx}^m), \tag{4.12}
\]

\[
\Sigma^X = - \sum_{i,k,l,m,n} (\epsilon_m + \epsilon_n) t_{mn}^{kl} X_{kl}^{ij} (2t_{mxj} - t_{jx}^m), \tag{4.13}
\]

with

\[
V_{mn}^{ij} = \langle ij | r_{12}^{-1} \hat{Q}_{12} f_{12} | mn \rangle, \tag{4.14}
\]

\[
B_{mn}^{ij} = \langle ij | f_{12} \hat{Q}_{12} (\hat{F}_1 + \hat{F}_2) \hat{Q}_{12} f_{12} | mn \rangle, \tag{4.15}
\]

\[
X_{mn}^{ij} = \langle ij | f_{12} \hat{Q}_{12} f_{12} | mn \rangle, \tag{4.16}
\]

where \( f_{12} \) is the geminal (an explicit function of \( r_{12} \)) and \( \hat{F}_n \) is the Fock operator of electron \( n \), i.e.,

\[
\hat{F}_n = \hat{T}_n + \hat{V}_n + \hat{J}_n - \hat{K}_n. \tag{4.17}
\]
\[ \hat{Q}_{12} = (1 - \hat{Q}_1)(1 - \hat{Q}_2) - \hat{V}_1 \hat{V}_2, \quad (4.18) \]

with

\[
\hat{Q}_1 = \sum_i^{\text{occ.}} \varphi_i(r_1) \int dr_3 \varphi_i^*(r_3) \hat{P}_{31}, \quad (4.19)
\]

\[
\hat{V}_1 = \sum_a^{\text{vir.}} \varphi_a(r_1) \int dr_3 \varphi_a^*(r_3) \hat{P}_{31}, \quad (4.20)
\]

where \( \hat{P}_{31} \) permutes electronic coordinates 1 and 3. We utilize the SP ansatz, the Generalized Brillouin Condition (GBC) and Extended Brillouin Condition (EBC).\textsuperscript{21,50,77} Due to the GBC and EBC approximations, as with MP2-F12, the X term cancels with terms within B and we are left with only 2 terms

\[
\Sigma_{x}^{\text{GF2-F12}} = 2 \Sigma_x^V + \Sigma_x^{BX}, \quad (4.21)
\]

\[
\Sigma_x^V = \sum_j^{\text{occ.}} \frac{5}{8} V_{xj}^j - \frac{1}{8} V_{jx}^x, \quad (4.22)
\]

\[
\Sigma_x^{BX} = \sum_j^{\text{occ.}} \frac{7}{32} (BX)_{xj}^j + \frac{1}{32} (BX)_{jx}^j, \quad (4.23)
\]

\[
V_{ij}^{ij} = \langle ij | r_{12}^{ij} \hat{Q}_{12} f_{12} | mn \rangle, \quad (4.24)
\]

\[
(BX)_{ij}^{ij} = \langle ij | f_{12} \hat{Q}_{12} [ \hat{F}_1 + \hat{F}_2, f_{12} ] | mn \rangle. \quad (4.25)
\]

Unfortunately, due to the decoupling of the orbital summations when transitioning over to the MC integrable expressions, the resulting integrals that are expanded are \textit{not} symmetric with respect to exchange of orbitals \( x \) and \( j \). To derive the correct correction for the \( x \)th orbital the following equations must be used

\[
\Sigma_x^V = \sum_j^{\text{occ.}} \frac{5}{16} \left( V_{xj}^j + V_{jx}^x \right) - \frac{1}{16} \left( V_{xj}^j + V_{jx}^x \right), \quad (4.26)
\]

\[
\Sigma_x^{BX} = \sum_j^{\text{occ.}} \frac{7}{64} \left( (BX)_{xj}^j + (BX)_{jx}^j \right) + \frac{1}{64} \left( (BX)_{xj}^j + (BX)_{jx}^j \right). \quad (4.27)
\]
The partitions of the GF2-F12 energy $\Sigma^V_x$ and $\Sigma^{BX}_x$ after expanding the projector $\hat{Q}_{12}$ are

$$\Sigma^V_x = \frac{5}{16} \sum_j \langle x| r^{-1}_{12} \hat{Q}_{12} f_{12} |x\rangle + \frac{5}{16} \sum_j \langle j|x| r^{-1}_{12} \hat{Q}_{12} f_{12} |j\rangle - \frac{1}{16} \sum_j \langle x| r^{-1}_{12} \hat{Q}_{12} f_{12} |x\rangle - \frac{1}{16} \sum_j \langle j|x| r^{-1}_{12} \hat{Q}_{12} f_{12} |j\rangle,$$  \hspace{0.5cm} (4.28)

$$\Sigma^{BX}_x = \frac{7}{64} \sum_j \langle x| f_{12} \hat{Q}_{12} [\hat{F}_1 + \hat{F}_2, f_{12}] |x\rangle + \frac{7}{64} \sum_j \langle j|x| f_{12} \hat{Q}_{12} [\hat{F}_1 + \hat{F}_2, f_{12}] |j\rangle + \frac{1}{64} \sum_j \langle x| f_{12} \hat{Q}_{12} [\hat{F}_1 + \hat{F}_2, f_{12}] |x\rangle + \frac{1}{64} \sum_j \langle j|x| f_{12} \hat{Q}_{12} [\hat{F}_1 + \hat{F}_2, f_{12}] |j\rangle.$$  \hspace{0.5cm} (4.29)

It has been shown that at the $\gamma$ value that minimizes the MP2-F12 energy, the MP2-F12 correction reduces to the $V$ term. The same property holds for the GF2 correction or,

$$\Sigma^{GF2-F12}_x = \Sigma^V_x$$  \hspace{0.5cm} (4.30)

This is only satisfied at the minimum of $\Sigma^{GF2-F12}_x$ with respect to our chosen parameters if the shape of the correlation factor simulates the cusped hole pair function accurately, but from our previous MP2-F12 studies the $V$ term alone can be quite accurate. With Ten-no’s correlation factor, we set $\gamma = 1.1$ and investigate if $V$ still achieves chemical accuracy in the GF2-F12 context.

Due to the large number of equations formed from the Fock operator and $\hat{Q}_{12}$ for $\Sigma^{BX}_x$, it is left to Appendix B. The derivation for $\Sigma^V_x$ will illustrate the key concepts of MC-GF2-F12. Expanding $\hat{Q}_{12}$, i.e., substituting Eq. (4.18) into Eq. (4.28), we obtain

$$\Sigma^V_{F12} = \Sigma^V_{2e,x} + \Sigma^V_{3e,x} + \Sigma^V_{4e,x},$$  \hspace{0.5cm} (4.31)
with

\[
\Sigma_{2e,x}^V = \frac{5}{16} \sum_j \left( \langle xj \rvert f_{12} \rangle \frac{f_{12}}{r_{12}} \lvert xj \rangle + \langle jx \rvert f_{12} \rangle \frac{f_{12}}{r_{12}} \lvert jx \rangle \right) \\
- \frac{1}{16} \sum_j \left( \langle xj \rvert \frac{f_{12}}{r_{12}} \lvert jx \rangle + \langle jx \rvert \frac{f_{12}}{r_{12}} \lvert jx \rangle \right),
\]

(4.32)

\[
\Sigma_{3e,x}^V = -\frac{5}{8} \sum_{jk} \left( \langle xjk \rvert f_{23} \rangle \frac{f_{23}}{r_{12}} \lvert kxj \rangle + \langle jxk \rvert f_{23} \rangle \frac{f_{23}}{r_{12}} \lvert jxk \rangle \right) \\
+ \frac{1}{8} \sum_{j,k} \left( \langle xjk \rvert \frac{f_{23}}{r_{12}} \lvert kxj \rangle + \langle jxk \rvert \frac{f_{23}}{r_{12}} \lvert jxk \rangle \right),
\]

(4.33)

\[
\Sigma_{4e,x}^V = \frac{5}{16} \sum_{j,k,l} \left( \langle xkl \rvert f_{34} \rangle \frac{f_{34}}{r_{12}} \lvert klxj \rangle + \langle jxkl \rvert f_{34} \rangle \frac{f_{34}}{r_{12}} \lvert jxkl \rangle \right) \\
- \frac{1}{16} \sum_{j,k,l} \left( \langle xkl \rvert \frac{f_{34}}{r_{12}} \lvert klxj \rangle + \langle jxkl \rvert \frac{f_{34}}{r_{12}} \lvert jxkl \rangle \right) \\
- \frac{5}{16} \sum_{j} \sum_{a,b} \left( \langle xjab \rvert \frac{f_{34}}{r_{12}} \lvert abxj \rangle + \langle jxab \rvert \frac{f_{34}}{r_{12}} \lvert jxab \rangle \right) \\
+ \frac{1}{16} \sum_{j} \sum_{a,b} \left( \langle xjab \rvert \frac{f_{34}}{r_{12}} \lvert abxj \rangle + \langle jxab \rvert \frac{f_{34}}{r_{12}} \lvert jxab \rangle \right),
\]

(4.34)

where the subscripts on the left-hand sides indicate the number of electrons involved in the integral and the orbital, \(x\), being correlated with GF2-F12. These are then converted into MC integrable form by

\[
\Sigma_{2e,x}^V = \int d\mathbf{r}_1 d\mathbf{r}_2 F_{2e,x}^V(r_1, r_2),
\]

(4.35)

\[
\Sigma_{3e,x}^V = \int \int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 F_{3e,x}^V(r_1, r_2, r_3),
\]

(4.36)

\[
\Sigma_{4e,x}^V = \int \int \int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 d\mathbf{r}_4 F_{4e,x}^V(r_1, r_2, r_3, r_4),
\]

(4.37)
with

\[
F_{2e,\alpha}^{V}(r_1, r_2) = \frac{5}{16} \frac{f_{12}(X_{11}O_{22} + O_{11}X_{22})}{r_{12}} - \frac{1}{16} \frac{f_{12}(X_{12}O_{21} + O_{12}X_{21})}{r_{12}},
\]

(4.38)

\[
F_{3e,\alpha}^{V}(r_1, r_2, r_3) = -\frac{5}{8} \frac{f_{23}(X_{13}O_{22}O_{31} + O_{13}X_{22}O_{31})}{r_{12}} + \frac{1}{8} \frac{f_{23}(X_{12}O_{23}O_{31} + O_{12}X_{23}O_{31})}{r_{12}},
\]

(4.39)

\[
F_{4e,\alpha}^{V}(r_1, r_2, r_3, r_4) = \frac{5}{16} \frac{f_{34}(X_{13}O_{24}O_{31}O_{42} + O_{13}X_{24}O_{31}O_{42})}{r_{12}} - \frac{1}{16} \frac{f_{34}(X_{14}O_{23}O_{31}O_{42} + O_{14}X_{23}O_{31}O_{42})}{r_{12}}
\]

\[
- \frac{5}{16} \frac{f_{34}(X_{13}O_{24}V_{31}V_{42} + O_{13}X_{24}V_{31}V_{42})}{r_{12}} + \frac{1}{16} \frac{f_{34}(X_{14}O_{23}V_{31}V_{42} + O_{14}X_{23}V_{31}V_{42})}{r_{12}},
\]

(4.40)

where

\[
X_{pq} = \psi^*_p(r_p)\psi_q(r_q),
\]

(4.41)

\[
O_{pq} = \sum_{\text{occ.}} \psi^*_i(r_p)\psi_i(r_q),
\]

(4.42)

\[
V_{pq} = \sum_{\text{vir.}} \psi^*_a(r_p)\psi_a(r_q).
\]

(4.43)

### 4.3 Monte Carlo integration and redundant walker algorithm

Similarly to MP2-F12, each of the contributions to the GF2-F12 correction is evaluated by the Metropolis MC method utilizing the redundant-walker algorithm.\textsuperscript{55} Each MC step for the corrections is evaluated as a ratio between the integrand and a judiciously chosen weight function. The weight function must cancel the singularities as the integrand, be always positive, and should generally have the same maxima and tails of the integrand.
The $V$ term [Eq. (4.31)] is evaluated as

$$
\sum V_{2e,x} = \int \int dr_1 dr_2 F_{2e,x}^V(r_1, r_2)
\approx \frac{1}{N} \sum_{n=1}^{N} \frac{F_{2e}^V(r_1^{[n]}, r_2^{[n]})}{w_{2e}(r_1^{[n]}, r_2^{[n]})},
$$

(4.44)

$$
\sum V_{3e,x} = \int \int \int dr_1 dr_2 dr_3 F_{3e,x}^V(r_1, r_2, r_3)
\approx \frac{1}{N} \sum_{n=1}^{N} \frac{1}{m} \sum_{k=1}^{m} \frac{F_{3e,x}^V(r_1^{[n]}, r_2^{[n]}, r_3^{[n]})}{w_{2e}(r_1^{[n]}, r_2^{[n]})w_{1e}(r_3^{[n]})},
$$

(4.45)

$$
\sum V_{4e,x} = \int \int \int \int dr_1 dr_2 dr_3 dr_4 F_{4e,x}^V(r_1, r_2, r_3, r_4)
\approx \frac{1}{N} \sum_{n=1}^{N} \frac{2!}{m(m-1)} \sum_{k=1}^{m-1} \sum_{l=k+1}^{m} \frac{F_{4e,x}^V(r_1^{[n]}, r_2^{[n]}, r_3^{[n]}, r_4^{[n]})}{w_{2e}(r_1^{[n]}, r_2^{[n]})w_{1e}(r_3^{[n]}_k)w_{1e}(r_4^{[n]}_l)},
$$

(4.46)

where $N$ is the total number of MC steps and $m$ is the number of “redundant walkers” (see below for the definition). There are two different types of walkers used for these integrands, those with strongly correlated coordinates or ”pair walkers” whose weight function is given by,

$$
w_{2e}(r_1, r_2) = \frac{1}{N_{2e}} \frac{g(r_1)g(r_2)}{r_{12}},
$$

(4.47)

and one-electron ”walkers” for coordinates with no singularity with the weight function as,

$$
w_{1e}(r_3) = \frac{g(r_3)}{N_{1e}},
$$

(4.48)

where $N_{1e}$ and $N_{2e}$ are normalization coefficients determined analytically. The $g(r)$ function is chosen to be a sum of $n$ atom-centered $s$-type GTOs per atom,

$$
g(r) = \sum_{\alpha=1}^{n_{atom}} c_{A}^{(\alpha)} e^{-c_{A}^{(\alpha)} |r-r_{A}|^2}
$$

(4.49)

where $n_{atom}$ is the number of atoms, $r_A$ is the position of the $A$th atom, and $n_{\lambda}$ is number of gaussians used. Usually two gaussians are used, one tight and one diffuse to mimic the electron density. For the ionization potential of $C_{60}$ and $C_{70}$, a third is added to account for the large size of the molecular...
orbits towards the HOMO.

The reason for this separation of walker coordinates can be seen in the $F_{3e}^V$ [Eq. (4.39)] and $F_{4e}^V$ [Eq. (4.40)] integrands. Electrons 1 and 2 and connected through the $r_{12}^{-1}$ operator and thus require a pair walker to cancel the singularity. Electrons 3 and 4 have no such connection to any other coordinate and can therefore be sampled according to the one-electron walker distribution that mimics the electron density. The merit of the redundant walker algorithm is that to calculate the densities and weight functions for each walker (pair or one-electron) costs only $O(m)$ but as seen in Eq. (4.46), we get $m(m − 1)/2$ samples, increasing the sampling efficiency of our MC algorithm by a factor of $O(m)$. However, Eq. (4.44) and is unaffected by the algorithm, making the overall benefit of the redundant algorithm for MC-GF2-F12 difficult to predict.

4.4 Results

Calculations for the aromatic molecules and polythiophenes at their optimized geometries were performed with the MC-GF2 and MC-GF2-F12 methods. 512 redundant walkers were used for MC-GF2 and 40 redundant one-electron and two-electron walkers were used for MC-GF2-F12. The statistical uncertainty was taken after 7 blocking transformations and all corrections were calculated at the aug-cc-pVDZ basis set level with the frozen core approximation. The number of MC steps performed for each molecule are given in Table B.1 in Appendix B.2. For every 8 processors used for an MC calculation, one of them was a GPU running MC-GF2 and the other 7 were CPUs running MC-GF2-F12.

MC-GF2-F12 corrections to the first ionization potential are shown in Table 4.1. Corrections for all molecules, no matter how large, were converged to a statistical error of 0.03 eV or less except for the $2B_u$ state of H2TPP which was generally more difficult to converge. Results from the deterministic method at the aug-cc-pVTZ basis set level are compiled for comparison. Our method shows near perfect agreement with the deterministic method with an average error from the deterministic result of 0.049 eV for the MC-GF2-F12/V method and 0.018 eV for the MC-GF2-F12/VBX method. Although the statistical error for the MAE in the case of the VBX results is about the same as the actual error, the error is at most 0.07 eV. This proves again the result from the previous study, that results equivalent to aug-cc-pVTZ can be derived from aug-cc-pVDZ using the EBC approximation if the exchange term is retained and appropriate $\gamma$ is chosen.

Scaling of the statistical error with system size ($n$) of the MC-GF2-F12 correction for a single orbital is shown in figure 4.1. The empirical scaling is shown to be approximately linear with number of basis functions. The scaling per MC step was defined previously as $O(n^2)$. Combining
Table 4.1: First ionization potentials for aromatic molecules and polythiophenes (T1-T7) in eV. MC-GF2/MC-GF2-F12-V/VBX were done at the aug-cc-pVDZ level with $\gamma=1.1$ using the STG. GF2-F12 results from Ten-no using the aug-cc-pVTZ basis set are compared. Statistical errors are shown for MC-GF2/MC-GF2-F12-V/VBX results.

<table>
<thead>
<tr>
<th>Molecule</th>
<th>MC-GF2/MC-GF2-F12-V</th>
<th>MC-GF2/MC-GF2-F12-VBX</th>
<th>GF2-F12</th>
</tr>
</thead>
<tbody>
<tr>
<td>benzene</td>
<td>9.174 ± 0.001</td>
<td>9.217 ± 0.001</td>
<td>9.235</td>
</tr>
<tr>
<td>naphthalene</td>
<td>7.981 ± 0.002</td>
<td>8.025 ± 0.003</td>
<td>8.045</td>
</tr>
<tr>
<td>anthracene</td>
<td>7.231 ± 0.004</td>
<td>7.278 ± 0.005</td>
<td>7.289</td>
</tr>
<tr>
<td>pyrene</td>
<td>7.246 ± 0.005</td>
<td>7.291 ± 0.006</td>
<td>7.314</td>
</tr>
<tr>
<td>coronene</td>
<td>7.131 ± 0.01</td>
<td>7.177 ± 0.013</td>
<td>7.185</td>
</tr>
<tr>
<td>ovalene</td>
<td>6.496 ± 0.021</td>
<td>6.542 ± 0.026</td>
<td>6.563</td>
</tr>
<tr>
<td>H2P $<em>{2B</em>{3u}}$</td>
<td>6.671 ± 0.019</td>
<td>6.713 ± 0.022</td>
<td>6.718</td>
</tr>
<tr>
<td>H2P $_{2A_u}$</td>
<td>7.019 ± 0.009</td>
<td>7.066 ± 0.011</td>
<td>7.095</td>
</tr>
<tr>
<td>H2TPP $<em>{2B</em>{3u}}$</td>
<td>6.125 ± 0.027</td>
<td>6.172 ± 0.038</td>
<td>6.155</td>
</tr>
<tr>
<td>H2TPP $_{2A_u}$</td>
<td>6.653 ± 0.021</td>
<td>6.691 ± 0.03</td>
<td>6.754</td>
</tr>
<tr>
<td>T1</td>
<td>8.857 ± 0.002</td>
<td>8.891 ± 0.004</td>
<td>8.914</td>
</tr>
<tr>
<td>T2</td>
<td>7.72 ± 0.007</td>
<td>7.748 ± 0.011</td>
<td>7.747</td>
</tr>
<tr>
<td>T3</td>
<td>7.238 ± 0.012</td>
<td>7.269 ± 0.018</td>
<td>7.257</td>
</tr>
<tr>
<td>T4</td>
<td>7.001 ± 0.013</td>
<td>7.039 ± 0.021</td>
<td>7.026</td>
</tr>
<tr>
<td>T5</td>
<td>6.804 ± 0.016</td>
<td>6.84 ± 0.026</td>
<td>6.825</td>
</tr>
<tr>
<td>T6</td>
<td>6.732 ± 0.018</td>
<td>6.775 ± 0.029</td>
<td>6.767</td>
</tr>
<tr>
<td>T7</td>
<td>6.677 ± 0.019</td>
<td>6.715 ± 0.029</td>
<td>6.695</td>
</tr>
<tr>
<td>MAE</td>
<td>0.049 ± 0.012</td>
<td>0.018 ± 0.017</td>
<td>-</td>
</tr>
</tbody>
</table>

this with the empirical scaling using the formula $\sigma = \tau N^2$ where $\sigma$ is the statistical error, $\tau$ is the time per MC step, and $N$ is the number of MC steps to achieve a specific error ($O(n)$), the total method scales as $O(n^4)$ for correcting a single orbital. Size dependence of correlating all orbitals in the basis set has not yet been observed.

The GF2-F12 corrections for three different correlation factors were computed for water, methane, and benzene shown in Table 4.2. Moving from STG to YC to SJ shows increasing statistical error with the SJ having a much larger error than the other two. It is speculated that the derivatives of the YC and SJ factors present in the kinetic energy vastly differ from the weight function used to sample the integrals, which could lead to higher statistical errors from Ten-no to YC to SJ. Due to the complexity of the integrals and sheer number of them, it is difficult to justify this result. It is also unclear why the increase in statistical errors seems to manifest more prominently in the larger molecule, Benzene. It is also unknown why these sort of increased errors seem to appear in MC-GF2-F12 theory but not MC-MP2-F12 theory.

To test the scalability and accuracy of the MC-GF2-F12 method for large molecules, calcula-
tions for C\textsubscript{60} and C\textsubscript{70} were run at the aug-cc-pVDZ basis set level. The calculated IPs are shown in Table 4.3 and all errors are converged to 0.04 eV or less for both fullerenes, with 0.043 eV being chemical accuracy. The MC-GF2 method achieves 6.70 eV and 6.80 eV for the IPs of C\textsubscript{60} and C\textsubscript{70} having substantial errors of 0.84 and 0.70 eV from the experimental values. Adding in the F12 correction improves the IP for C\textsubscript{60} and C\textsubscript{70} by 0.47 eV and 0.45 eV respectively resulting in errors of 0.47 eV and 0.22 eV respectively from the experimental values. The MC-GF2-F12 methodology although substantially improving the MC-GF2 results with the aug-cc-pVDZ basis set suffers from a large error from experiment in the case of C\textsubscript{60} and a moderate error in the case of C\textsubscript{70}. There are many possible reasons for this such as basis-set used in the calculation, GBC and EBC approximations, lack of higher-orders of perturbation theory, and the diagonal and frequency-independent approximations used in deriving this methodology. More accuracy could be obtainable by extending our formalism to non-diagonal frequency-dependent Green’s function theory which is also able to derive electron affinities (EAs). We plan to also include MC-GF3 and MC-GF4 corrections in
Table 4.2: MC-GF2-F12 calculations for the IP of small molecules with Ten-no, YC, and SJ correlation factors using the $\gamma$ values of 1.1, 1.6 and 1.1 respectively. Values for the MC-GF2-F12(VBX) correction and statistical error were taken after $4.59 \times 10^7$.

<table>
<thead>
<tr>
<th>Correlation factor</th>
<th>H$_2$O</th>
<th>CH$_4$</th>
<th>Benzene</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ten-no’s STG</td>
<td>-0.578 ± 0.001</td>
<td>-0.331 ± 0.001</td>
<td>-0.402 ± 0.003</td>
</tr>
<tr>
<td>Yukawa-Coulomb</td>
<td>-0.569 ± 0.002</td>
<td>-0.327 ± 0.001</td>
<td>-0.411 ± 0.010</td>
</tr>
<tr>
<td>Slater-Jastrow</td>
<td>-0.569 ± 0.007</td>
<td>-0.335 ± 0.006</td>
<td>-0.384 ± 0.078</td>
</tr>
</tbody>
</table>

Table 4.3: Calculated ionization potentials of C$_{60}$ and C$_{70}$ at the aug-cc-pVDZ. For C$_{60}$ MC-GF2 and MC-GF2-F12 were run for $3.67 \times 10^8$ and $6.15 \times 10^7$ MC steps respectively. For C$_{70}$ MC-GF2 and MC-GF2-F12(VBX) were run for $3.14 \times 10^8$ and $4.54 \times 10^7$ MC steps respectively.

<table>
<thead>
<tr>
<th>Molecule</th>
<th>MC-GF2</th>
<th>MC-GF2-F12(VBX)</th>
<th>Exp.$^{1}$</th>
</tr>
</thead>
<tbody>
<tr>
<td>C$_{60}$</td>
<td>6.70 ± 0.02</td>
<td>7.17 ± 0.03</td>
<td>7.64 ± 0.02</td>
</tr>
<tr>
<td>C$_{70}$</td>
<td>6.80 ± 0.02</td>
<td>7.25 ± 0.04</td>
<td>7.47 ± 0.20</td>
</tr>
</tbody>
</table>

$^{1}$ Experimental reference values taken from Ref.$^{122}$.

The future.

4.5 Conclusions

We have implemented the first ever MC-GF2-F12 method scaling up to large numbers of processors. IPs of many small to medium-sized organic molecules were calculated and the accuracy of the method to the corresponding deterministic study was verified. The MAE of MC-GF2-F12(VBX) from the previous deterministic result with the aug-cc-pVTZ basis set was 0.049/0.018 eV using the aug-cc-pVDZ basis set for our calculations. The size dependence of the statistical error of the MC-GF2-F12 method was empirically measured to be $O(n^4)$ where $n$ is the number of basis functions for correlating a single orbital. Unlike MC-MP2-F12, strange behavior was observed for the statistical error when using different correlation factors. The reason has not yet been found as to why MC calculations with some factors are more difficult to converge.

MC-GF2-F12 calculations for the IP of C$_{60}$ and C$_{70}$ were shown to be tractable on the Blue Waters supercomputer. MC-GF2 alone suffered from large errors nearing 1.0 eV with the F12 contribution improving each IP by around 0.5 eV. This shows the necessity of this correction to derive agreement with experiment. C$_{60}$ suffered from a large error of 0.47 eV from the experimental value while C$_{70}$ had a much smaller error of 0.22 eV from experiment. The causes of the larger molecule, C$_{70}$, having much higher accuracy to the experimental value is unknown. In the future, results from MC-GF3 and MC-GF4 will be added and a more thorough study of the advantages and
disadvantages will be conducted. The IPs of two fullerenes were calculated here, but the electron attachment (EA) energies are a much more sought after quantity in the case of the fullerenes. Another future study will also be conducted adding in frequency dependence and non-diagonal terms into the self-energy expression to further improve the accuracy, and the applicability of the method to correlated EAs.
Chapter 5: Conclusions

In Chapter 2, a stochastic algorithm for evaluating the MP2-F12 energies, called MC-MP2-F12, for use on the Blue Waters Supercomputer was developed. Scaling with respect to the number of processors was found to be almost exactly linear and scaling with system size to achieve a relative error was found to be $O(n^4)$, one order lower scaling than the corresponding deterministic method. In Chapter 3, the MC-MP2-F12 methodology developed in Chapter 2 was utilized to test 17 different correlation factors, many of which had never been tested before with the MP2-F12 methodology. All of the highly performing factors had very similar shape from 0 to about 1.5 Bohr. Contrarily, the long-range behavior of the correlation seems to not affect the energetics at all, except possibly in the case of the linear $r_{12}$ factor. In Chapter 4, the MC-F12 techniques were extended to calculate ionization potentials stochastically evaluating the second-order self-energy. The MC-GF2 and MC-GF2-F12 methods were combined for a fully stochastic calculation of IPs. The same set of molecules used in the recent deterministic GF2-F12 paper were studied, and mean average error of 0.018 eV was achieved from the deterministic results with the MC-GF2-F12/VBX method. System-size scaling was shown to scale as $O(n^4)$ for correlating a single orbital. As a test of practical scalability, calculations were performed for the MC-GF2-F12 corrected IPs of C$_{60}$ and C$_{70}$. With just the aug-cc-pVDZ basis set, errors of 0.47 and 0.22 eV from experiment were achieved for C$_{60}$ and C$_{70}$ respectively. The explanation as to why higher accuracy to experiment is derived for C$_{70}$ is unknown.

The novel MC-F12 methodology has proven to be an efficient way to get MP2 and GF2 corrections near the CBS limit with one-order lower scaling in system size than the corresponding deterministic methods with near 100% scaling efficiency with number of CPUs. Increasing the applicability and accuracy for large molecules such as the fullerenes tested with C$_{60}$ and C$_{70}$ is the end goal of the development of these methods. The first step to attempting to derive more accuracy is combining the existing MC-GF2-F12 with MC-GF3 and MC-GF4. An additional factor is the MC-GF2-F12 method presented here is diagonal and frequency independent. By relaxing these approximations, reduction of the large error in C$_{60}$ may be possible. Deriving the MC-GF2-F12 formalism straight from the propagator, as was done in a recent study$^{30}$ would offer the ability to calculate electron affinities for band gaps of solids.

A different avenue not studied here, is solving the relativistic time-independent Schrödinger equation by deriving relativistic MC-MP2 and MC-MP2-F12. This presents new and interesting
challenges such as a generally complex basis set, spin-dependence in the Hamiltonian as well as more complicated spin dependence in the basis set. If enough reduction of scaling is achieved, heavy metal complexes could be benchmarked with high parallel efficiency. If this is successful, similar generalizations to Green’s function theory might also be possible.

Some holes exist in the comprised studies. Chapter 3 contains a hole in whether or not the cusp condition being satisfied at \( r_{12} = 0 \) is quantitatively important. The fact that one can expand the Slater-type geminal by a set of Gaussians suggests that it may not be that important, but a definitively tested answer in the MP2-F12 context still remains to be seen. Also, the simplest Slater-Jastrow based on a small number of parameters was used (so as to see what the optimal parameter values were) where a more optimized one with more parameters may have given better performance. And lastly, the analysis of the long-range behavior of the correlation factor was derived from CH\(_4\) and H\(_2\)O systems. These small molecules give limited insight into the behavior at large \( r_{12} \) and larger molecules need to be tested to see if the trend holds in general. The empirical scaling observations have their highest contribution from a calculation with about 1400 basis functions so it would also be interesting to see if the behavior observed for the statistical error maintains its \( O(n) \) form or higher order dependence in \( n \) which had a small prefactor arises. Lastly, the reason as to why some correlation factors are more difficult to converge and what features determine their statistical error is still a mystery. If some level of knowledge here was obtainable, a correlation factor may be able to be developed which gives both high recovery of the CBS limit as well as minimization of the variance. Insights into this area could also be extended to QMC and other \textit{ab initio} methods simulated by MC.
References


[113] The number of redundant walkers in Ref. [56] also referred to the numbers of one-electron and two-electron walkers.


Appendix A: MC-MP2-F12 details

A.1 Derivations of Eqs. (2.69), (2.70), and (2.78)

We show that the terms carrying the numerical factor of $7/32$ in Eqs. (2.69) and (2.70) arise from the first term of Eq. (2.56). The other terms in Eqs. (2.69) and (2.70) as well as higher-order integrands are derived similarly.

Substituting Eq. (2.72) into the first term of Eq. (2.56), we obtain

\[
\frac{7}{32} \sum_{i,j} \langle ij | \hat{T}_1 + \hat{T}_2, f_{12} | ij \rangle = \frac{7}{32} \sum_{i,j} \langle ij | f_{12} \left\{ \frac{f_{12}^{(a)}}{r_{12}} + f_{12}^{(b)} \right\} | ij \rangle \\
- \frac{7}{32} \sum_{i,j} \langle ij | f_{12} \left\{ \frac{f_{12}^{(c)}}{r_{12}} + f_{12}^{(d)} \right\} | r_{12} \cdot (\nabla_1 - \nabla_2) | ij \rangle
\] (A.1)

\[
= \frac{7}{32} \int dr_1 dr_2 \sum_{i,j} \varphi_i^*(r_1) \varphi_j^*(r_2) \left\{ \frac{f_{12}^{(a)}}{r_{12}} + f_{12}^{(b)} \right\} \varphi_j(r_1) \varphi_j(r_2) \\
- \frac{7}{32} \int dr_1 dr_2 \sum_{i,j} \varphi_i^*(r_1) \varphi_j^*(r_2) \left\{ \frac{f_{12}^{(c)}}{r_{12}} + f_{12}^{(d)} \right\} \varphi_j(r_2) \cdot \nabla_1 \varphi(r_1) \\
+ \frac{7}{32} \int dr_1 dr_2 \sum_{i,j} \varphi_i^*(r_1) \varphi_j^*(r_2) \left\{ \frac{f_{12}^{(c)}}{r_{12}} + f_{12}^{(d)} \right\} \varphi_j(r_1) \cdot \nabla_2 \varphi_j(r_2) \\
= \frac{7}{32} \int dr_1 dr_2 \frac{f_{12}^{(a)} O_{11} O_{22}}{r_{12}} + \frac{7}{32} \int dr_1 dr_2 \frac{f_{12}^{(b)} O_{11} O_{22}}{r_{12}} \\
+ \frac{7}{32} \int dr_1 dr_2 \frac{f_{12}^{(c)} (O_{11} O_{22}' - O_{11}' O_{22})}{r_{12}} \\
+ \frac{7}{32} \int dr_1 dr_2 \frac{f_{12}^{(d)} (O_{11} O_{22}' - O_{11}' O_{22})}{r_{12}},
\] (A.4)

where a change in the order of summations and integrations is made in Eq. (A.4), an essential step for bringing the expression into a MC-integrable form.

Next, we show that the first term of Eq. (2.78) with the numerical factor of $7/16$ comes from the first term of Eq. (2.59).
Substituting the definition of the exchange operator,

\[ \hat{K}_1 = \sum_k^{\text{occ.}} \int d r_3 \varphi_k^*(r_3) \frac{1}{r_{13}} \hat{P}_{13} \varphi_k(r_3), \]  

(A.5)

into the \( \hat{K}_1 \) contribution of the first term of Eq. (2.59), we obtain the following. If

\[ M = -\frac{7}{32} \sum_{i,j}^{\text{occ.}} \langle i| f_{12} \left[ \hat{K}_1, f_{12} \right]|j \rangle \]

then

\[ M = -\frac{7}{32} \sum_{i,j}^{\text{occ.}} \int d r_1 d r_2 \varphi_i^*(r_1) \varphi_j^*(r_2) f_{12} \]

\[ \times \sum_k^{\text{occ.}} \int d r_3 \varphi_k^*(r_3) \frac{1}{r_{13}} \hat{P}_{13} \varphi_k(r_3) f_{12} \varphi_i(r_1) \varphi_j(r_2) \]

\[ + \frac{7}{32} \sum_{i,j}^{\text{occ.}} \int d r_1 d r_2 \varphi_i^*(r_1) \varphi_j^*(r_2) f_{12} f_{12} \]

\[ \times \sum_k^{\text{occ.}} \int d r_3 \varphi_k^*(r_3) \frac{1}{r_{13}} \hat{P}_{13} \varphi_k(r_3) \varphi_i(r_1) \varphi_j(r_2) \]  

(A.6)

\[ = -\frac{7}{32} \int d r_1 d r_2 d r_3 \sum_{i,j,k}^{\text{occ.}} \varphi_i^*(r_1) \varphi_j^*(r_2) f_{12} \varphi_k^*(r_3) \frac{1}{r_{13}} \varphi_k(r_3) f_{32} \varphi_i(r_1) \varphi_j(r_2) \]

\[ + \frac{7}{32} \int d r_1 d r_2 d r_3 \sum_{i,j,k}^{\text{occ.}} \varphi_i^*(r_1) \varphi_j^*(r_2) f_{12} f_{12} \varphi_k^*(r_3) \frac{1}{r_{13}} \varphi_k(r_3) \varphi_i(r_1) \varphi_j(r_2) \]  

(A.7)

\[ = -\frac{7}{32} \int d r_1 d r_2 d r_3 O_{13} O_{22} O_{31} f_{12} f_{32} \]

\[ + \frac{7}{32} \int d r_1 d r_2 d r_3 O_{13} O_{22} O_{31} f_{12} f_{12} \]  

(A.8)

\[ = -\frac{7}{32} \int d r_1 d r_2 d r_3 O_{12} O_{32} O_{21} f_{13} f_{23} \]

\[ + \frac{7}{32} \int d r_1 d r_2 d r_3 O_{12} O_{32} O_{21} f_{12} f_{23} \]  

(A.9)

in the last step of which, a coordinate interchange of \( r_2 \leftrightarrow r_3 \) (additionally \( r_1 \leftrightarrow r_2 \) in the last term) is carried out to ensure that the singularity occurs only at \( r_{12} = 0 \), which is also essential for bringing the final expression into a form most convenient for MC integration. Furthermore, using a coordinate interchange of \( r_1 \leftrightarrow r_2 \), we can immediately show that the \( \hat{K}_2 \) contribution of the first
term of Eq. (2.59) is

\[ -\frac{7}{32} \sum_{i,j}^{\text{occ.}} \langle ij | f_{12} [\hat{K}_2, f_{12}] | ij \rangle = -\frac{7}{32} \sum_{i,j}^{\text{occ.}} \langle ij | f_{12} [\hat{K}_1, f_{12}] | ij \rangle, \tag{A.10} \]

yielding

\[ -\frac{7}{32} \sum_{i,j}^{\text{occ.}} \langle ij | f_{12} [\hat{K}_1 + \hat{K}_2, f_{12}] | ij \rangle = \frac{7}{16} \int \int \int dr_1 dr_2 dr_3 \frac{f_{23}(f_{23} - f_{13})}{r_{12}} O_{12} O_{33} O_{21}. \tag{A.11} \]

The other term of Eq. (2.59) and higher-rank integrands can be derived similarly.

### A.2 Weight functions

Two types of weight function are used in MC-MP2-F12 calculations. One is the one-electron weight function, \( w_{1e}(r) \), in the form of Eq. (2.86), and the other is the two-electron weight function, \( w_{2e}(r_1, r_2) \), of Eq. (2.84). They both are defined in terms of a linear combination, \( g(r) \), of two atom-centered \( s \)-type GTOs [Eq. (2.85)]. For each atom, the two exponents and expansion coefficients in \( g(r) \) are given in Tables A.1 and A.2. The only important consideration in defining \( g(r) \) is that it should be more diffuse than the most diffuse of the integrands, so that the Metropolis algorithm over-samples rather than under-samples. An incidence of under-sampling can be detected as a vertical jump in the evolution of the statistical uncertainty.

Table A.1: The weight function [Eq. (2.85)] for DZ and ADZ.

<table>
<thead>
<tr>
<th>Atom (A)</th>
<th>( c_A^{(1)} )</th>
<th>( c_A^{(2)} )</th>
<th>( \xi_A^{(1)} )</th>
<th>( \xi_A^{(2)} )</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>0.5</td>
<td>0.6</td>
<td>0.05</td>
<td>0.15</td>
</tr>
<tr>
<td>C</td>
<td>1.0</td>
<td>1.0</td>
<td>0.10</td>
<td>0.25</td>
</tr>
<tr>
<td>N</td>
<td>2.5</td>
<td>1.4</td>
<td>0.25</td>
<td>0.30</td>
</tr>
<tr>
<td>O</td>
<td>3.0</td>
<td>1.8</td>
<td>0.30</td>
<td>0.37</td>
</tr>
<tr>
<td>F</td>
<td>4.5</td>
<td>1.8</td>
<td>0.45</td>
<td>0.35</td>
</tr>
</tbody>
</table>
Table A.2: The weight function \([\text{Eq. (2.85)}]\) for TZ and ATZ.

<table>
<thead>
<tr>
<th>Atom (A)</th>
<th>(c_A^{(1)})</th>
<th>(\xi_A^{(1)})</th>
<th>(c_A^{(2)})</th>
<th>(\xi_A^{(2)})</th>
</tr>
</thead>
<tbody>
<tr>
<td>H</td>
<td>0.5</td>
<td>0.6</td>
<td>0.05</td>
<td>0.10</td>
</tr>
<tr>
<td>C</td>
<td>1.0</td>
<td>0.8</td>
<td>0.10</td>
<td>0.13</td>
</tr>
<tr>
<td>N</td>
<td>2.5</td>
<td>1.0</td>
<td>0.25</td>
<td>0.19</td>
</tr>
<tr>
<td>O</td>
<td>3.0</td>
<td>1.0</td>
<td>0.30</td>
<td>0.22</td>
</tr>
<tr>
<td>F</td>
<td>4.5</td>
<td>1.2</td>
<td>0.45</td>
<td>0.28</td>
</tr>
</tbody>
</table>
Appendix B: MC-GF2-F12 details

This Appendix contains the detailed derivation of the $\Sigma_{\chi}^{BX}$ term and all equations therein and a reference table of all MC steps used to perform the calculations in Table 4.1.

B.1 $\Sigma_{\chi}^{BX}$ full derivation

Inserting Eq. (4.18) into equation Eq. (4.29) results in five different parts,

$$\Sigma_{\chi}^{BX} = \Sigma_{2e,\chi}^{T} + \Sigma_{3e,\chi}^{T} + \Sigma_{4e,\chi}^{T} + \Sigma_{5e,\chi}^{K} + \Sigma_{5e,\chi}^{K}.$$  \hspace{1cm} (B.1)

The kinetic and exchange terms are

$$\Sigma_{2e,\chi}^{T} = \frac{7}{64} \sum_{j} \langle x| f_{12} \hat{T}_{2} + \hat{T}_{2}, f_{12} | j \rangle + \frac{7}{64} \sum_{j} \langle j| f_{12} \hat{T}_{2} + \hat{T}_{2}, f_{12} | j \rangle$$

$$+ \frac{1}{64} \sum_{j} \langle x| f_{12} \hat{T}_{2} + \hat{T}_{2}, f_{12} | j \rangle + \frac{1}{64} \sum_{j} \langle j| f_{12} \hat{T}_{2} + \hat{T}_{2}, f_{12} | x \rangle,$$  \hspace{1cm} (B.2)

$$\Sigma_{3e,\chi}^{T} = -\frac{7}{32} \sum_{jk} \langle xkj| f_{12} \hat{T}_{2} + \hat{T}_{2}, f_{23} | kxj \rangle$$

$$-\frac{7}{32} \sum_{jk} \langle jkx| f_{12} \hat{T}_{2} + \hat{T}_{2}, f_{23} | jxk \rangle$$

$$-\frac{1}{32} \sum_{jk} \langle xkj| f_{12} \hat{T}_{2} + \hat{T}_{2}, f_{23} | kxj \rangle$$

$$-\frac{1}{32} \sum_{jk} \langle jkx| f_{12} \hat{T}_{2} + \hat{T}_{2}, f_{23} | kxj \rangle,$$  \hspace{1cm} (B.3)
\[ \Sigma_{4e,x}^T = \frac{7}{64} \sum_{j,k,l}^{\text{occ.}} \langle x j k l | f_{12} [\hat{T}_3 + \hat{T}_4, f_{34}] | k l x j \rangle \\
+ \frac{7}{64} \sum_{j,k,l}^{\text{occ.}} \langle j x k l | f_{12} [\hat{T}_3 + \hat{T}_4, f_{34}] | k l j x \rangle \\
+ \frac{1}{64} \sum_{j,k,l}^{\text{occ.}} \langle x j k l | f_{12} [\hat{T}_3 + \hat{T}_4, f_{34}] | k l j x \rangle \\
+ \frac{1}{64} \sum_{j,k,l}^{\text{occ.}} \langle j x k l | f_{12} [\hat{T}_3 + \hat{T}_4, f_{34}] | k l x j \rangle \\
- \frac{7}{64} \sum_{j}^{\text{vir.}} \sum_{a,b}^{\text{occ.}} \langle x j a b | f_{12} [\hat{T}_3 + \hat{T}_4, f_{34}] | a b x j \rangle \\
- \frac{7}{64} \sum_{j}^{\text{vir.}} \sum_{a,b}^{\text{occ.}} \langle j x a b | f_{12} [\hat{T}_3 + \hat{T}_4, f_{34}] | a b j x \rangle \\
- \frac{1}{64} \sum_{j}^{\text{vir.}} \sum_{a,b}^{\text{occ.}} \langle x j a b | f_{12} [\hat{T}_3 + \hat{T}_4, f_{34}] | a b j x \rangle \\
- \frac{1}{64} \sum_{j}^{\text{vir.}} \sum_{a,b}^{\text{occ.}} \langle j x a b | f_{12} [\hat{T}_3 + \hat{T}_4, f_{34}] | a b x j \rangle, \quad (B.4) \]

and

\[ \Sigma_{3e,x}^K = -\frac{7}{64} \sum_{j}^{\text{occ.}} \langle x j | f_{12} [\hat{K}_1 + \hat{K}_2, f_{12}] | x j \rangle \\
- \frac{7}{64} \sum_{j}^{\text{occ.}} \langle j x | f_{12} [\hat{K}_1 + \hat{K}_2, f_{12}] | j x \rangle \\
- \frac{1}{64} \sum_{j}^{\text{occ.}} \langle x j | f_{12} [\hat{K}_1 + \hat{K}_2, f_{12}] | j x \rangle \\
- \frac{1}{64} \sum_{j}^{\text{occ.}} \langle j x | f_{12} [\hat{K}_1 + \hat{K}_2, f_{12}] | x j \rangle, \quad (B.5) \]
\[ \Sigma_{4e,x}^K = \frac{7}{32} \sum_{jk}^{\text{occ.}} \langle jx| f_{12} [\hat{K}_2 + \hat{K}_3, f_{23}] |jkx \rangle \\
+ \frac{7}{32} \sum_{jk}^{\text{occ.}} \langle jx| f_{12} [\hat{K}_2 + \hat{K}_3, f_{23}] |kxj \rangle \\
+ \frac{1}{32} \sum_{jk}^{\text{occ.}} \langle jx| f_{12} [\hat{K}_2 + \hat{K}_3, f_{23}] |kxj \rangle \\
+ \frac{1}{32} \sum_{jk}^{\text{occ.}} \langle jx| f_{12} [\hat{K}_2 + \hat{K}_3, f_{23}] |jkx \rangle , \quad (B.6) \]

\[ \Sigma_{5e,x}^K = -\frac{7}{64} \sum_{jkl}^{\text{occ.}} \langle jxkl| f_{12} [\hat{K}_3 + \hat{K}_4, f_{34}] |klxj \rangle \\
- \frac{7}{64} \sum_{jkl}^{\text{occ.}} \langle jxkl| f_{12} [\hat{K}_3 + \hat{K}_4, f_{34}] |lxjk \rangle \\
- \frac{1}{64} \sum_{jkl}^{\text{occ.}} \langle jxkl| f_{12} [\hat{K}_3 + \hat{K}_4, f_{34}] |lxjk \rangle \\
- \frac{1}{64} \sum_{jkl}^{\text{occ.}} \langle jxkl| f_{12} [\hat{K}_3 + \hat{K}_4, f_{34}] |kxjl \rangle \\
+ \frac{7}{64} \sum_{j} \sum_{a,b}^{\text{vir.}} \langle jab| f_{12} [\hat{K}_3 + \hat{K}_4, f_{34}] |abxj \rangle \\
+ \frac{7}{64} \sum_{j} \sum_{a,b}^{\text{vir.}} \langle jxab| f_{12} [\hat{K}_3 + \hat{K}_4, f_{34}] |abxj \rangle \\
+ \frac{1}{64} \sum_{j} \sum_{a,b}^{\text{vir.}} \langle jab| f_{12} [\hat{K}_3 + \hat{K}_4, f_{34}] |labxj \rangle \\
+ \frac{1}{64} \sum_{j} \sum_{a,b}^{\text{vir.}} \langle jxab| f_{12} [\hat{K}_3 + \hat{K}_4, f_{34}] |labxj \rangle , \quad (B.7) \]

where we use the fact that only the kinetic-energy (\(\hat{T}_n\)) and exchange (\(\hat{K}_n\)) operators in the Fock operator [Eq. (4.17)] do not commute with \(f_{12}\) and, therefore,

\[ [\hat{F}_1 + \hat{F}_2, f_{12}] = [\hat{T}_1 + \hat{T}_2, f_{12}] - [\hat{K}_1 + \hat{K}_2, f_{12}] . \quad (B.8) \]

Note that the dimension in of the exchange terms in [Eqs. (B.5)–(B.7)] is one-electron higher.
than the number in the bra and ket. This is due to the exchange operator containing a one-electron electron projector like $\hat{Q}_{12}$.

The MC-GF2-F12 equations have the same properties as the MC-MP2-F12 equations with the additional complication of symmeterization and the $X_{pq}$ orbital products. Following these steps we arrive at the following MC-integrable expressions:

\[
\Sigma^T_{2e,x} = \int \int dr_1 dr_2 F^T_{2e,x}(r_1, r_2) \\
+ \int \int dr_1 dr_2 F^T_{2e,s}(r_1, r_2), 
\]  
(B.9)

\[
\Sigma^T_{3e,x} = \int \int \int dr_1 dr_2 dr_3 F^T_{3e,x}(r_1, r_2, r_3) \\
+ \int \int \int dr_1 dr_2 dr_3 F^T_{3e,s}(r_1, r_2, r_3), 
\]  
(B.10)

\[
\Sigma^T_{4e,x} = \int \int \int \int dr_1 dr_2 dr_3 dr_4 F^T_{4e,x}(r_1, r_2, r_3, r_4) \\
+ \int \int \int \int dr_1 dr_2 dr_3 dr_4 F^T_{4e,s}(r_1, r_2, r_3, r_4), 
\]  
(B.11)

\[
\Sigma^K_{3e,x} = \int \int \int dr_1 dr_2 dr_3 F^K_{3e,x}(r_1, r_2, r_3), 
\]  
(B.12)

\[
\Sigma^K_{4e,x} = \int \int \int \int dr_1 dr_2 dr_3 dr_4 F^K_{4e,x}(r_1, r_2, r_3, r_4), 
\]  
(B.13)

\[
\Sigma^K_{5e,x} = \int \cdots \int dr_1 \cdots dr_5 F^K_{5e,x}(r_1, r_2, r_3, r_4, r_5), 
\]  
(B.14)
Expanding out the operators in Eqs. (B.2)-(B.7) we arrive at:

\[ F_{2e,x}^{T_1}(r_1, r_2) = \frac{7}{64} f_{12} f_{12}^{(a)} (X_{11} O_{22} + O_{11} X_{22}) \]
\[ + \frac{7}{64} f_{12} f_{12}^{(c)} (X_{11} O_{22} + O_{11} X_{22} - X'_{11} O_{22} - O'_{11} X_{22}) \]
\[ + \frac{1}{64} f_{12} f_{12}^{(a)} (X_{12} O_{21} + O_{12} X_{21}) \]
\[ + \frac{1}{64} f_{12} f_{12}^{(c)} (X_{12} O_{21} + O_{12} X_{21} - X'_{12} O_{21} - O'_{12} X_{21}) \]  

(B.15)

\[ F_{2e,x}^{T_2}(r_1, r_2) = \frac{7}{64} f_{12} f_{12}^{(b)} (X_{11} O_{22} + O_{11} X_{22}) \]
\[ + \frac{7}{64} f_{12} f_{12}^{(d)} (X_{11} O_{22} + O_{11} X_{22} - X'_{11} O_{22} - O'_{11} X_{22}) \]
\[ + \frac{1}{64} f_{12} f_{12}^{(b)} (X_{12} O_{21} + O_{12} X_{21}) \]
\[ + \frac{1}{64} f_{12} f_{12}^{(d)} (X_{12} O_{21} + O_{12} X_{21} - X'_{12} O_{21} - O'_{12} X_{21}) \]  

(B.16)

\[ F_{3e,x}^{T_1}(r_1, r_2, r_3) = -\frac{7}{32} f_{23} f_{12}^{(a)} (X_{31} O_{22} + O_{31} X_{22}) O_{13} \]
\[ - \frac{7}{32} f_{23} f_{12}^{(c)} (X_{22} O_{31} + O'_{22} X_{31} - X'_{31} O_{22} + O'_{31} X_{22}) O_{13} \]
\[ - \frac{1}{32} f_{23} f_{12}^{(a)} (X_{32} O_{21} + O_{32} X_{21}) O_{13} \]
\[ - \frac{1}{32} f_{23} f_{12}^{(c)} (X_{32} O_{21} + O'_{32} X_{21} - X'_{21} O_{32} - O'_{21} X_{32}) O_{13} \]  

(B.17)

\[ F_{3e,x}^{T_2}(r_1, r_2, r_3) = -\frac{7}{32} f_{23} f_{12}^{(b)} (X_{31} O_{22} + O_{31} X_{22}) O_{13} \]
\[ - \frac{7}{32} f_{23} f_{12}^{(d)} (X_{22} O_{31} + O'_{22} X_{31} - X'_{31} O_{22} - O'_{31} X_{22}) O_{13} \]
\[ - \frac{1}{32} f_{23} f_{12}^{(b)} (X_{32} O_{21} + O_{32} X_{21}) O_{13} \]
\[ - \frac{1}{32} f_{23} f_{12}^{(d)} (X_{32} O_{21} + O'_{32} X_{21} - X'_{21} O_{32} - O'_{21} X_{32}) O_{13} \]  

(B.18)
\[ F_{4e,1}^{T1}(r_1, r_2, r_3, r_4) \]
\[
= \frac{7}{64} \left( f_{34} f_{12}^{(a)}(X_{31} O_{42} + O_{31} X_{42})O_{13} O_{24} \right)
\]
\[
+ \frac{7}{64} \left( f_{34} f_{12}^{(c)}(X_{42} O_{31} + O_{42} X_{31} - X_{31} O_{42} - O_{31} X_{42})O_{13} O_{24} \right)
\]
\[
+ \frac{1}{64} \left( f_{34} f_{12}^{(a)}(X_{32} O_{41} + O_{32} X_{41})O_{13} O_{24} \right)
\]
\[
+ \frac{1}{64} \left( f_{34} f_{12}^{(c)}(X_{32} O_{41} + O_{32} X_{41} - X_{31} O_{42} - O_{31} X_{42})O_{13} O_{24} \right)
\]
\[
- \frac{7}{64} \left( f_{34} f_{12}^{(a)}(X_{31} O_{42} + O_{31} X_{42})V_{13} V_{24} \right)
\]
\[
+ \frac{7}{64} \left( f_{34} f_{12}^{(c)}(X_{42} O_{31} + O_{42} X_{31} - X_{31} O_{42} - O_{31} X_{42})V_{13} V_{24} \right)
\]
\[
+ \frac{1}{64} \left( f_{34} f_{12}^{(a)}(X_{32} O_{41} + O_{32} X_{41})V_{13} V_{24} \right)
\]
\[
+ \frac{1}{64} \left( f_{34} f_{12}^{(c)}(X_{32} O_{41} + O_{32} X_{41} - X_{31} O_{42} - O_{31} X_{42})V_{13} V_{24} \right),
\]
\[
(B.19)
\]
\[
F_{4e,2}^{T2}(r_1, r_2, r_3, r_4)
\]
\[
= \frac{7}{64} \left( f_{34} f_{12}^{(b)}(X_{31} O_{42} + O_{31} X_{42})O_{13} O_{24} \right)
\]
\[
+ \frac{7}{64} \left( f_{34} f_{12}^{(d)}(X_{42} O_{31} + O_{42} X_{31} - X_{31} O_{42} - O_{31} X_{42})O_{13} O_{24} \right)
\]
\[
+ \frac{1}{64} \left( f_{34} f_{12}^{(b)}(X_{32} O_{41} + O_{32} X_{41})O_{13} O_{24} \right)
\]
\[
+ \frac{1}{64} \left( f_{34} f_{12}^{(d)}(X_{32} O_{41} + O_{32} X_{41} - X_{41} O_{32} - O_{41} X_{32})O_{13} O_{24} \right)
\]
\[
- \frac{7}{64} \left( f_{34} f_{12}^{(b)}(X_{31} O_{42} + O_{31} X_{42})V_{13} V_{24} \right)
\]
\[
+ \frac{7}{64} \left( f_{34} f_{12}^{(d)}(X_{42} O_{31} + O_{42} X_{31} - X_{31} O_{42} - O_{31} X_{42})V_{13} V_{24} \right)
\]
\[
+ \frac{1}{64} \left( f_{34} f_{12}^{(b)}(X_{32} O_{41} + O_{32} X_{41})V_{13} V_{24} \right)
\]
\[
+ \frac{1}{64} \left( f_{34} f_{12}^{(d)}(X_{32} O_{41} + O_{32} X_{41} - X_{41} O_{32} - O_{41} X_{32})V_{13} V_{24} \right),
\]
\[
(B.20)
\]
for the kinetic contributions and

\[
F_{3e,x}^K(r_1, r_2, r_3) = \frac{7}{32} f_{23}(f_{23} - f_{13}) \left( X_{12} O_{33} + O_{12} X_{33} \right) O_{21} \\
+ \frac{1}{32} f_{23}(f_{23} - f_{13}) \left( X_{13} O_{32} + O_{13} X_{32} \right) O_{21},
\]

(B.21)

\[
F_{4e,x}^K(r_1, r_2, r_3, r_4) = -\frac{7}{32} f_{24}(f_{23} - f_{13}) \left( X_{43} O_{21} + O_{43} X_{21} \right) O_{34} O_{12} \\
- \frac{1}{32} f_{24}(f_{23} - f_{13}) \left( X_{41} O_{23} + O_{41} X_{23} \right) O_{34} O_{12} \\
- \frac{1}{32} f_{24}(f_{23} - f_{13}) \left( X_{42} O_{21} + O_{42} X_{21} \right) O_{34} O_{12} \\
- \frac{1}{32} f_{24}(f_{23} - f_{13}) \left( X_{42} O_{21} + O_{42} X_{21} \right) O_{34} O_{12},
\]

(B.22)

\[
F_{5e,x}^K(r_1, r_2, r_3, r_4, r_5) = \frac{7}{32} f_{35}(f_{14} - f_{24}) \left( X_{32} O_{54} + O_{32} X_{54} \right) O_{31} O_{45} O_{21} \\
+ \frac{1}{32} f_{35}(f_{14} - f_{24}) \left( X_{34} O_{52} + O_{34} X_{52} \right) O_{31} O_{45} O_{21} \\
- \frac{7}{32} f_{35}(f_{14} - f_{24}) \left( X_{32} O_{54} + O_{32} X_{54} \right) V_{31} V_{45} O_{21} \\
- \frac{1}{32} f_{35}(f_{14} - f_{24}) \left( X_{34} O_{52} + O_{34} X_{52} \right) V_{31} V_{45} O_{21}.
\]

(B.23)

for the exchange contributions with

\[
X_{pq}' = \varphi^*_x(r_p) r_{12} \cdot \nabla_q \varphi_x(r_q),
\]

(B.24)

\[
O_{pq}' = \sum_{\text{occ}} \varphi^*_i(r_p) r_{12} \cdot \nabla_q \varphi_i(r_q).
\]

(B.25)

See Appendix A from the previous MC-MP2-F12 study for a detailed derivation of equations analogous to Eqs. (B.15) and (B.16).\(^5\) The integrand of Eq. (B.9) is divided into two terms, \(F^{T_1}_{2e,x}\) and \(F^{T_2}_{2e,x}\) resulting from the first having a singularity at \(r_{12} = 0\) and the other not, thus requiring different weight functions in MC integrations. The subscripts of \(r_{12}\) in Eq. (B.25) are always set to ‘12’ to ensure the pair walker for all integrals samples \(r_1\) and \(r_2\). The factors of the geminal, \(f^{(a)}\), \(f^{(b)}\), \(f^{(c)}\), and \(f^{(d)}\), are defined as

\[
\hat{T}_1 + \hat{T}_2, f_{12} = \frac{f^{(a)}_{12}}{r_{12}} + f^{(b)}_{12} - \left( \frac{f^{(c)}_{12}}{r_{12}} + f^{(d)}_{12} \right) r_{12} \cdot (\nabla_1 - \nabla_2),
\]

(B.26)
for any geminal. The expressions in Eqs. (B.15)-(B.23) are evaluated as

\[ \sum_{T_{2e,x}} \approx \frac{1}{N} \sum_{n=1}^{N} \frac{F_{T_{2e,x}}^{1} (r^{[n]}_1, r^{[n]}_2)}{w_{2e}(r^{[n]}_1, r^{[n]}_2)} \]

\[ + \frac{1}{N} \sum_{n=1}^{N} \frac{2!}{m(m-1)} \sum_{k=1}^{m-1} \sum_{l=k+1}^{m} \frac{F_{T_{2e,x}}^{2} (r^{[n]}_1, r^{[n]}_2)}{w_{1e}(r^{[n]}_1)w_{1e}(r^{[n]}_2)}, \]

\( \text{(B.27)} \)

\[ \sum_{T_{3e,x}} \approx \frac{1}{N} \sum_{n=1}^{N} \frac{2!}{m(m-1)(m-2)} \]

\[ \times \sum_{k=1}^{m-2} \sum_{l=k+1}^{m-1} \sum_{h=l+1}^{m} \frac{F_{T_{3e,x}}^{1} (r^{[n]}_1, r^{[n]}_2, r^{[n]}_3)}{w_{2e}(r^{[n]}_1, r^{[n]}_2)w_{1e}(r^{[n]}_3)w_{1e}(r^{[n]}_4)}, \]

\[ \text{(B.28)} \]

\[ \sum_{T_{4e,x}} \approx \frac{1}{N} \sum_{n=1}^{N} \frac{2!}{m(m-1)} \]

\[ \times \sum_{k=1}^{m-1} \sum_{l=k+1}^{m} \sum_{h=k+1}^{m} \sum_{i=l+1}^{m} \frac{F_{T_{4e,x}}^{1} (r^{[n]}_1, r^{[n]}_2, r^{[n]}_3, r^{[n]}_4)}{w_{2e}(r^{[n]}_1, r^{[n]}_2)w_{2e}(r^{[n]}_3, r^{[n]}_4)w_{1e}(r^{[n]}_1)w_{1e}(r^{[n]}_2)} \]

\[ \text{(B.29)} \]

\[ \sum_{K_{3e,x}} \approx \frac{1}{N} \sum_{n=1}^{N} \frac{2!}{m(m-1)} \]

\[ \times \sum_{k=1}^{m-1} \sum_{l=k+1}^{m} \frac{F_{K_{3e,x}}^{1} (r^{[n]}_1, r^{[n]}_2, r^{[n]}_3)}{w_{2e}(r^{[n]}_1, r^{[n]}_2)w_{1e}(r^{[n]}_3)}, \]

\[ \text{(B.30)} \]

\[ \sum_{K_{4e,x}} \approx \frac{1}{N} \sum_{n=1}^{N} \frac{2!}{m(m-1)} \]

\[ \times \sum_{k=1}^{m-1} \sum_{l=k+1}^{m} \frac{F_{K_{4e,x}}^{1} (r^{[n]}_1, r^{[n]}_2, r^{[n]}_3, r^{[n]}_4)}{w_{2e}(r^{[n]}_1, r^{[n]}_2)w_{2e}(r^{[n]}_3, r^{[n]}_4)w_{1e}(r^{[n]}_1)w_{1e}(r^{[n]}_2)}, \]

\[ \text{(B.31)} \]
where weight functions \( w_{2e} \) and \( w_{1e} \) here are identical to Eqs. (4.47) and (4.48), respectively, and thus the same set of pair and one-electron walkers used for the \( V \) term are used for \( BX \) as well as for the MC-GF2 procedure. As with the \( V \) term, the sampling efficiency of the first term of Eq. (B.27) is not affected by the redundant-walker algorithm. Contrarily the sampling efficiency of the second term is increased by a factor of \( O(m) \). The higher dimensional integrals have an effect on the overall efficiency that is difficult to determine.

### B.2 MC steps used

The number of MC steps used for each molecule is given in the table below. Calculations for all molecules were calculated utilizing the Blue Waters supercomputer.

Table B.1: Number of MC steps used for each molecule in billions.

<table>
<thead>
<tr>
<th>Molecule</th>
<th>MC-GF2 steps</th>
<th>MC-GF2-F12 steps</th>
</tr>
</thead>
<tbody>
<tr>
<td>benzene</td>
<td>0.51</td>
<td>0.43</td>
</tr>
<tr>
<td>naphthalene</td>
<td>0.41</td>
<td>0.33</td>
</tr>
<tr>
<td>anthracene</td>
<td>0.33</td>
<td>0.22</td>
</tr>
<tr>
<td>pyrene</td>
<td>0.31</td>
<td>0.19</td>
</tr>
<tr>
<td>coronene</td>
<td>0.21</td>
<td>0.08</td>
</tr>
<tr>
<td>ovalene</td>
<td>0.21</td>
<td>0.06</td>
</tr>
<tr>
<td>H2P(^2)B(_3)u)</td>
<td>0.20</td>
<td>0.07</td>
</tr>
<tr>
<td>H2P(^2)A(_u)</td>
<td>0.39</td>
<td>0.13</td>
</tr>
<tr>
<td>H2TPP(^2)B(_u)</td>
<td>1.48</td>
<td>0.23</td>
</tr>
<tr>
<td>H2TPP(^2)A(_u)</td>
<td>1.86</td>
<td>0.29</td>
</tr>
<tr>
<td>T1</td>
<td>0.57</td>
<td>0.45</td>
</tr>
<tr>
<td>T2</td>
<td>0.42</td>
<td>0.33</td>
</tr>
<tr>
<td>T3</td>
<td>0.33</td>
<td>0.19</td>
</tr>
<tr>
<td>T4</td>
<td>0.96</td>
<td>0.44</td>
</tr>
<tr>
<td>T5</td>
<td>0.96</td>
<td>0.3</td>
</tr>
<tr>
<td>T6</td>
<td>2.0</td>
<td>0.53</td>
</tr>
<tr>
<td>T7</td>
<td>1.81</td>
<td>0.39</td>
</tr>
</tbody>
</table>