Quantum Monte Carlo:

Some Theoretical and Numerical Studies.

by

Heather Louise Gordon, B.Sc. Chem.

A Thesis

submitted to the Department of Chemistry

in partial fulfilment of the requirements

for the degree of

Master of Science

July 1983

Brock University

St. Catharines, Ontario

Abstract

In Part I, theoretical derivations for Variational Monte Carlo calculations are compared with results from a numerical calculation of He; both indicate that minimization of the ratio estimate of E_{var}, denoted $\hat{\tilde{E}}_{MC}$, provides different optimal variational parameters than does minimization of the variance of $\hat{\tilde{E}}_{MC}$. Similar derivations for Diffusion Monte Carlo calculations provide a theoretical justification for empirical observations made by other workers.

In Part II, Importance sampling in prolate spheroidal coordinates allows Monte Carlo calculations to be made of E_{var} for the vdW molecule He_2, using a simplifying partitioning of the Hamiltonian and both an HF-SCF and an explicitly correlated wavefunction. Improvements are suggested which would permit the extension of the computational precision to the point where an estimate of the interaction energy could be made.

Acknowledgements

Table of Contents, continued...

List of Tables

List of Figures

I. Introduction

The Schroedinger Equation

$$H\Psi = E\Psi \qquad (1.1)$$

is exactly solvable for one-electron atoms and molecules; for many-electron atoms and molecules, approximate methods must be employed since the interelectronic repulsion terms r_{ij}^{-1} in the Hamiltonian H causes the Schroedinger Equation to be inseparable in any coordinate system. The Variational Method provides an approximate solution for the ground-state energy of a system without solving the Schroedinger Equation.

The Variational Theorem [1] states that, given a trial wavefunction ψ_t which is well-behaved, normalizable and obeys the boundary conditions of the system, then

$$E_o \leqq E_{var} = \int \psi_t^* H \psi_t \, d\tau \bigg/ \int \psi_t^* \psi_t \, d\tau \qquad (1.2)$$

where E_o is the lowest eigenvalue of H and E_{var} is the so-called variational energy. As the trial wavefunction more closely approximates the true exact wavefunction, ψ_o, the variational energy approaches E_o. Thus the variational integral (1.2) provides an upper bound for E_o.

The mathematical form that ψ_t can assume varies widely: a compact, physically meaningful wavefunction may give rise to intractable integrals while one which tries to avoid such difficulties may need so many terms and parameters that all physical meaning is lost. An example of the former is an explicitly correlated wavefunction, so-called because it involves expansions in terms of the interelectronic coordinates. The presence of such terms efficiently incorporates the effect of electron correlation into the energy calculation but leads to very complicated integrals. Monte Carlo techniques for integration [2-5] can be used to numerically evaluate such integrals: the mean and variance of integrand values evaluated at a large number of representative points over the coordinate domain are calculated. The mean is an estimate of the value of the integral and the statistical variance is a measure of the precision of the estimate.

In practice, ψ_t is a function of a set of adjustable parameters, $\underline{\beta}$, which are to be optimized. In the Monte Carlo scheme, this has been done either by minimizing the Monte Carlo estimate $\hat{E}_{MC}$ of E_{var} (1.2) [6] or by minimizing the variance of $\hat{E}_{MC}$ [7]. The former is justified by the variational theorem while the latter has a statistical justification: a smaller variance means a more precise estimate. It can be shown that for ψ_o the statistical variance of the Monte Carlo energy estimate is zero [7, 8].

The objective of the first part of this work is to derive expressions in order to determine theoretically which optimization scheme is to be preferred given a Variational Monte Carlo calculation of E_{var}

(1.2). These results are compared with those found empirically from a Monte Carlo simulation of a simple, known system (Helium atom). The conclusions reached are then of use in the second part of the work, where the more complicated system of the van der Waals He_2 molecule is considered.

For He_2, Self-Consistent-Field (SCF) theory fails to predict the existence of a shallow minimum in the intermolecular energy curve because it neglects the effect of electron correlation [9]. Hence the use of explicitly correlated wavefunctions promises to be of use in _ab initio_ calculations of the interaction energy for such rare-gas systems [7, 10, 11].

In the second part of this thesis, the following are accomplished;

(i)   a Monte Carlo sampling procedure for the He_2 system is constructed;

(ii)   equations are developed which enable SCF results to be combined with Monte Carlo estimates of the effect of the inclusion of a correlation function with the SCF wavefunction;

(iii)   a computer program incorporating the above is written and different correlation functions are examined.

The integral in (1.2) is also a specific example of a quantum mechanical expectation value. Given an operator, O, the expectation or average value is given by

$$\langle O \rangle = \int \psi^*(\underline{x}) O(\underline{x}) \psi(\underline{x}) \, d\underline{x} \, \bigg/ \int \psi^*(\underline{x}) \psi(\underline{x}) \, d\underline{x} \tag{2.1}$$

given a distribution function $\psi^*(\underline{x}) \psi(\underline{x}) \, d\underline{x}$. Hence the E_{var} is the expectation value of the Hamiltonian, given the normalized everywhere non-negative, probability density $\rho(\underline{x})$

$$\rho(\underline{x}) = \psi_t^*(\underline{x}) \psi_t(\underline{x}) \, \bigg/ \int \psi_t^*(\underline{x}) \psi_t(\underline{x}) \, d\underline{x} \tag{2.2}$$

(See (2.5) below.)

The Monte Carlo Method offers a numerical solution to integrals of the type shown in (2.1) where the multidimensionality and inseparability of the particle coordinates make the analytic evaluations intractable. In particular, as mentioned in the introduction, this occurs when the integrand is a function of r_{ij}.

If O does not involve any differential operators, then (2.1) can be rearranged as

$$\langle O \rangle = \int O(\underline{x}) \rho(\underline{x}) \, d\underline{x} \quad . \tag{2.3}$$

It is shown in Appendix A that the Central Limit Theorem of probability gives

$$\langle \mathcal{O} \rangle \ = \ \lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} \mathcal{O}(\underset{\sim}{x}_i) \tag{2.4}$$

where $(\underset{\sim}{x}_i)$ are a set of points in configuration space, drawn from (2.2). In the case where $\mathcal{O}$ does involve differential operators, (2.1) can be written as

$$\langle \mathcal{O} \rangle \ = \ \int \left[\mathcal{O}(\underset{\sim}{x}) \psi_t(\underset{\sim}{x}) / \psi_t(\underset{\sim}{x}) \right] \rho(\underset{\sim}{x}) \, d\underset{\sim}{x} \tag{2.5}$$

and hence

$$\langle \mathcal{O} \rangle \ = \ \lim_{n\to\infty} n^{-1} \sum_{i=1}^{n} \left[\mathcal{O}(\underset{\sim}{x}_i) \psi_t(\underset{\sim}{x}_i) / \psi_t(\underset{\sim}{x}_i) \right] \ . \tag{2.6}$$

Equations (2.4) and (2.6) show that the integrals in (2.3) and (2.5) may be estimated by an average over random values of the function $\mathcal{O}$ or $(\mathcal{O}\psi_t/\psi_t)$ evaluated at n representative points $\underset{\sim}{x}_i$ from the continuous, normalized and non-negative probability function $\rho(\underset{\sim}{x})$. In the limit as the number of values included in the sum becomes infinite, the estimator becomes exact; the Monte Carlo estimate of $\langle \mathcal{O} \rangle$, where $n < \infty$ will be designated $\hat{\mathcal{O}}_{MC}$.

In addition to calculating $\hat{\mathcal{O}}_{MC}$, it is of importance to have some estimate of the variance of $\hat{\mathcal{O}}_{MC}$, which is associated with the precision of the Monte Carlo estimate of $\langle \mathcal{O} \rangle$. In Appendix A, it is shown that, for _independent_ samples, confidence intervals around $\hat{\mathcal{O}}_{MC}$ can be constructed from the estimate of the variance of $\hat{\mathcal{O}}_{MC}$, $\hat{Var}(\hat{\mathcal{O}}_{MC})$, the true

variance being $Var(\hat{O}_{MC})$. This information allows the theoretician to state with some degree of confidence that $\langle O \rangle$ lies within some interval surrounding $\hat{O}_{MC}$. Even if confidence levels cannot be established, as is the case with the use of correlated samples or of samples having unknown distributions, estimates of precision are useful, since it is desirable to consider both accuracy and precision as indicators of a successful "experiment" [24]. In the event that true expectation values are not known, as is the usual case in Quantum Chemical problems, some knowledge of the precision of estimates is especially important as accuracies cannot be established.

The normalization condition of the probability density $\rho(\underline{x})$ in the denominator of (2.1) presents the first obstacle in the computation of $\hat{O}_{MC}$. Quite often, the choice of $\psi_t(\underline{x})$ precludes the analytic evaluation of the normalization constant. In order to circumvent this problem, one of two approaches may be taken. The most common, as applied to Variational methods, is to use the Metropolis Monte Carlo technique [13]. In this method, the set of points $\underline{x}_i$ is the result of a random walk through configuration space formed by a Markov process [3]. For Metropolis sampling, the configurations $\underline{x}_k$ and $\underline{x}_{k+1}$ are not statistically independent. The advantage of the Metropolis scheme is that it is only necessary to evaluate ratios of $\rho(\underline{x})$ for successive configurations, thus the normalization integrals cancel out. The disadvantage is that correlation in the sample increases the variance and hence causes slow convergence towards the expectation value. Binder [21] points out the many considerations necessary for good Metropolis Monte Carlo work.

An alternative approach is to simulate both integrals in (2.1) using a random sample of configuration points, thus avoiding correlation in the sample. The disadvantage of this method lies in the expense of generating the random configuration points from the inversion of a probability density, as compared to the acception-rejection method employed by the Metropolis technique. However, the statistical theory for treating such samples is well-known; the straightforward analysis of variance is an advantage because the convergence of the Monte Carlo estimate towards the true answer is slow (generally on the order of $n^{-\frac{1}{2}}$, n being the sample size), and because the true answer is not generally known.

This second approach, which will be designated as Variational Monte Carlo (VMC), is used in this work. Recently a third approach called Diffusion Monte Carlo (DMC), has been applied by several workers [25-27]. DMC contrasts with VMC and Metropolis Monte Carlo in having the potential to estimate E_o rather than E_{var}. It is of interest to compare the statistics for DMC with those of VMC and a short chapter, V, is devoted to this purpose.

III. Principles of Monte Carlo Evaluation of Integrals

The evaluation of multiple integrals by Monte Carlo Methods is an extension of the one-dimensional case which will be examined for simplicity. Consider then, the integral I:

$$I = \int_a^b f(x)ds \qquad (3.1)$$

where $f(x)$ is continuous over the interval $a \leq x \leq b$ and assume that $f \epsilon L^2(a,b)$ (that is,

$$\int_a^b [f(x)]^2 dx$$

exists and hence I exists). Equation (3.1) can be expressed in terms of an expectation value

$$I \equiv \langle f/\rho \rangle = \int_a^b [f(x)/\rho(x)]\rho(x)dx \qquad (3.2)$$

over a normalized, continuous and non-negative density function, $\rho(x)$, defined in the region $a \leq x \leq b$:

$$\rho(x) \geq 0 , \quad a \leq x \leq b \qquad (3.3a)$$

$$1 = \int_a^b \rho(x)dx \qquad . \qquad (3.3b)$$

The $\langle f/\rho \rangle$ can be termed the expected value of a continuous random function. The variance of f/ρ, $\text{Var}(f/\rho)$ is defined by:

$$\text{Var}(f/\rho) = \langle (f/\rho)^2 \rangle - \langle f/\rho \rangle^2$$

$$= \int_a^b [f(x)/\rho(x)]^2 \rho(x) dx - \left[\int_a^b [f(x)/\rho(x)] \rho(x) dx \right]^2 . \quad (3.4)$$

The sample-mean Monte Carlo Method [19] numerically evaluates I using (3.2) by generating n random points $x_1, \ldots, x_n$ distributed according to $\rho(x)$ and estimating $\langle f/\rho \rangle$ as the mean of the $f(x_i)/\rho(x_i)$:

$$\hat{I} = n^{-1} \sum_{i=1}^n f(x_i)/\rho(x_i) . \quad (3.5)$$

$\hat{I}$ is an unbiased estimator and converges to the exact value I as $n \longrightarrow \infty$. The estimate of the variance of $\hat{I}$ is, from (3.4)

$$\hat{\text{Var}}(\hat{I}) = (n-1)^{-1} \sum_{i=1}^n \left([f(x_i)/\rho(x_i)]^2 \right) - n(n-1)^{-1} \hat{I}^2 . \quad (3.6)$$

For a sufficiently large sample, $n \cong (n-1)$ and

$$\hat{\text{Var}}(\hat{I}) = n^{-1} \sum_{i=1}^n [f(x_i)/\rho(x_i)]^2 - \hat{I}^2 . \quad (3.7)$$

Given that $\{x_i\}$ are n independent identically distributed random variables, then the $\{f(x_i)/\rho(x_i)\}$ are likewise distributed. Hence,

$$\langle f(x_i)/\rho(x_i)\rangle = I \quad , \quad i = 1, \ldots, n \tag{3.8a}$$

and

$$\mathrm{Var}\big(f(x_i)/\rho(x_i)\big) = \mathrm{Var}(I) \quad , \quad i = 1, \ldots, n \tag{3.8b}$$

where the angular brackets denote the expected value.

The expected value of the sum of the random variates is treated as follows:

$$\left\langle \sum_{i=1}^{n} f(x_i)/\rho(x_i)\right\rangle = \sum_{i=1}^{n} \langle f(x_i)/\rho(x_i)\rangle = nI \tag{3.9a}$$

and the variance of the sum is then

$$\mathrm{Var}\Big(\sum_{i=1}^{n} f(x_i)/\rho(x_i)\Big) = \sum_{i=1}^{n} \mathrm{Var}\big(f(x_i)/\rho(x_i)\big) = n\mathrm{Var}(I) \quad . \tag{3.9b}$$

The Central Limit Theorem (CLT) (see Appendix A) of probability states that such a sum of random variables is asymptotically normally distributed with mean nI and variance nVar(I):

$$p\Big(\sum_{i=1}^{n} f(x_i)/\rho(x_i)\Big) \sim N\big(nI, \, n\mathrm{Var}(I)\big) \quad , \quad n \longrightarrow \infty \tag{3.10}$$

and so the sample mean is also asymptotically normally distributed

$$p(I) \sim N\big(I, \, \mathrm{Var}(I)\big) \quad , \qquad n \longrightarrow \infty \quad . \tag{3.11}$$

The significance of this theorem is that it is possible to construct confidence intervals about $\hat{I}$ because the normal distribution has a known mathematical form. Hence it is possible to show that the probability of finding the true mean within 3 standard deviations of $\hat{I}$ is about 99.7%:

$$P\left(\hat{I} - 3\text{Var}^{\frac{1}{2}}(\hat{I}) < I < \hat{I} + 3\text{Var}^{\frac{1}{2}}(\hat{I})\right) \cong 0.997 \tag{3.12}$$

In particular, using $\widehat{\text{Var}}(\hat{I})$ (3.7), it can be stated that the true value I lies within the interval $\hat{I} \pm 3\widehat{\text{Var}}^{\frac{1}{2}}(\hat{I})$ approximately 99.7% of the time due to chance, given a sufficiently large sample.

The Monte Carlo variance (3.7) is proportional to n^{-1} and the standard deviation $\widehat{\text{Var}}^{\frac{1}{2}}(\hat{I})$ to $n^{-\frac{1}{2}}$. The statistical error decreases as the sample size increases, however the convergence is slow: in order to gain one significant digit in the precision of $\hat{I}$, the sample must be increased 100-fold. In practice, almost all Monte Carlo work employs some kind of and/or combination of variance-reduction techniques [19,20]. Importance sampling [28] is one such technique universally employed in all types of Quantum Monte Carlo Methods.

Consider the case where $\rho(x)$ is a uniform distribution between $a \leq x \leq b$; given the constraints (3.3) then

$$\rho(x) = (b - a)^{-1} \qquad a \leq x \leq b \quad . \tag{3.13}$$

Then from (3.2) and (3.5)

$$I = (b - a)<f>$$
(3.14)

and the Monte Carlo estimate is

$$\hat{I} = n^{-1}(b - a)\sum_{i=1}^{n} f(x_i) \qquad .$$
(3.15)

The variance of $\hat{I}$ is

$$Var(\hat{I}) = Var\left(n^{-1}(b - a)\sum_{i=1}^{n} f(x_i)\right)$$

$$= n^{-1}(b - a)^2 Var(f)$$

$$= n^{-1}(b - a)^2 \left[\int_a^b f^2(x)(b - a)^{-1}dx - \left(\int_a^b f(x)dx\right)^2\right]$$

$$= n^{-1}\left[(b - a)\int_a^b f^2(x) - I^2\right] \qquad .$$
(3.16)

An estimate of $\hat{Var}(\hat{I})$ is then, from (3.7),

$$\hat{Var}(\hat{I}) = n^{-1}(b - a)^2 \sum_{i=1}^{n} f^2(x_i) - \hat{I}^2$$
(3.17)

Equations (3.15) and (3.17) are the Crude Monte Carlo estimate and variance respectively; each element in the Monte Carlo sum has an identical weight. Since a configuration point x_j is just as likely to be chosen as a point x_k from the distribution in (3.13), even if

$f(x_j) \ll f(x_k)$, a great many sample points must be taken in order for points where $f(x)$ is large to dominate the Monte Carlo sum. Hence the variance of the Crude Monte Carlo estimate is large.

Importance sampling uses a suitably chosen $\rho(x)$ so that sample points are concentrated in areas where $f(x)$ is large; bias is removed by a suitable weighting scheme. Equation (3.2) shows that any suitable $\rho(x)$ complying with the constraints (3.3) will provide an estimate of I. It can be proven [19] that the minimum variance is equal to

$$\mathrm{Var}(f/\rho_o) = \left[\int_a^b |f(x)| dx \right]^2 - I^2 \qquad (3.18)$$

and occurs when

$$\rho_o(x) = |f(x)| \left/ \int_a^b |f(x)| dx \right. \qquad (3.19)$$

In other words, if $f(x)$ is everywhere non-negative, choosing $\rho(x)$ proportional to $f(x)$ gives a zero variance. In reality, since (3.1) is not known, $\rho_o(x)$ (3.19) will not be known. However, this result shows that if $\rho(x)$ is chosen to have a shape similar to $f(x)$, then the ratio f/ρ will be relatively constant and the variance of the estimate will be reduced. The optimal choice of ρ for a particular f, given the constraints (3.3), is a difficult analytical problem [19] and in general, is taken on the basis of knowledge about f and the ease of normalization.

Given an importance sampling density, $\rho(x)$, it is necessary to generate a set of random variables $\{x_i\}$ to use in (3.5) and (3.7).

Several procedures are available [29]; the inverse transform or probability
inversion method [29, 30] is used here.

Computer-generated pseudorandom numbers (so-called because they
are not truly random but comply with statistical tests of "randomness")
are drawn from standard distributions. Consider then a set of pseudorandom
numbers $\{P_i\}$ drawn from a univariate uniform distribution: P_i represents
the cumulative probability of $\rho(x)$ at $x = x_i$

$$P_i = \int_a^{x=x_i} \rho(x)\,dx \qquad .$$
$$\text{(3.20)}$$

The integral in (3.20) is performed and $x = x_i$ is solved for, given P_i.
Sometimes this is a difficult problem; in other cases more efficient methods
of generating $\{x_i\}$ are available [29].

The multivariate case, where $\rho = \rho(x_1, \ldots, x_r)$ is somewhat
similar; the following equations apply whether or not the variables
$x_1, \ldots, x_r$ are independent. Define reduced density functions

$$\rho_k(x_1, \ldots, x_k) \equiv \int \ldots \int \rho(x_1, \ldots, x_r)\,dx_{k+1}\ldots dx_r$$
$$k = 1, \ldots, r-1 \qquad \text{(3.21)}$$

and conditional density functions

$$\Gamma_{k+1}(x_{k+1}|x_1, \ldots, x_k) \equiv \rho_{k+1}(x_1, \ldots, x_{k+1})/\rho_k(x_1, \ldots, x_k) \qquad .$$
$$\text{(3.22)}$$

Since

$$\rho_1(x_1) \prod_{i=1}^{r-1} \Gamma(x_{i+1} | x_1, \ldots, x_i) = \rho(x_1, \ldots, x_r) \quad , \tag{3.23}$$

the r-dimensional random numbers can be generated by first selecting x_1 from $\rho_1(x_1)$ and then x_i, $i = 2, \ldots, r$ from $\Gamma_i(x_i | x_1, \ldots, x_{i-1})$ successively. There are r! ordered combinations to represent $x_1, \ldots, x_r$ and hence r! possibilities to generate the sets of random variables. For example, if $r = 2$, then

$$\rho(x_1, x_2) = \begin{cases} \rho_1(x_1) \Gamma_2(x_2 | x_1) \\ \rho_1(x_2) \Gamma_2(x_1 | x_2) \end{cases} \tag{3.24}$$

which promises two courses of action:

(1) find x_1 and then x_2 conditional on x_1, or

(2) find x_2 and then x_1 conditional on x_2.

If the variables are independent, all r! combinations are equivalent; otherwise the combination chosen depends on the efficiency of the inversion problems posed and on the how the variables are interdependent [29].

IV. Variational Monte Carlo: Theoretical Derivations and Parameter
Optimization [31]

In this chapter, formulae are derived for E_{var} and the variance of
the VMC energy estimate $Var(\hat{E}_{MC})$ in terms of smallness parameters, which
relate to the accuracy of ψ_t. The effect of minimizing either the E_{var} or
$Var(\hat{E}_{MC})$ with respect to variational parameters $\underset{\sim}{\beta}$ in ψ_t is examined.

Consider an unnormalized wavefunction, ψ_t. Expand ψ_t in terms of
the exact (normalized) ground state, ϕ_o, and (normalized) excited states,
ϕ_k, with the same symmetry as ϕ_o:

$$\psi_t = \phi_o + \sum_{k=1}^{\infty} \lambda_k \phi_k \tag{4.1}$$

where λ_k are smallness parameters [32] such that $\lambda_k \longrightarrow 0$ as $\psi_t \longrightarrow \phi_o$.
(In all equations that follow, the summation is assumed to run over all
excited states.) The variational energy E_{var} (1.2) is a quotient of two
expectation values:

$$E_{var} = \langle H \rangle / \langle S \rangle \tag{4.2}$$

where

$$H \equiv \psi_t^* H \psi_t \tag{4.3}$$

and

$$S \equiv \psi_t^* \psi_t \tag{4.4}$$

where H and S are functions of $R^{(i)} \equiv (\underline{r}_1^{(i)}, \ldots, \underline{r}_N^{(i)})$, the i-th member of a set of n points in the configuration space of N electrons.

Upon substituting (4.1) into (4.3) and (4.4) one obtains

$$\langle H \rangle = E_o + \sum_k \lambda_k^2 E_k \tag{4.5}$$

where E_o and E_k are the exact ground and k-th excited state energies, respectively, and similarly,

$$\langle S \rangle = 1 + \sum_k \lambda_k^2 \quad . \tag{4.6}$$

Thus,

$$E_{var} = E_o + \sum_k \lambda_k^2 (E_k - E_o) + O(\lambda^4) \quad ; \tag{4.7}$$

the well-known result that the variational energy is an upper bound to E_o and is of second order in a smallness parameter.

Assume a simple random sample of n configurations, $R^{(i)}$. The (crude) VMC estimate of the quotient

$$\hat{E}_{MC} = n^{-1} \sum_i H(R^{(i)}) \bigg/ n^{-1} \sum_i S(R^{(i)}) \tag{4.8}$$

converges to E_{var} (4.7) as $n \longrightarrow \infty$ (a "consistent" estimate is provided). Statisticians refer to (4.8) as a ratio estimate. It has a negligible bias for large n [33]:

$$\langle \hat{E}_{MC} \rangle = E_{var} + O(n^{-1}), \tag{4.9}$$

where E_{var} is given by (4.7). The distribution of (4.8) is asymptotically normal, subject to mild conditions on the distributions of H and S [34].

The variance of the ratio estimator (4.8) is given by [33]:

$$Var(\hat{E}_{MC}) = (n\langle S \rangle^2)^{-1}\Big(Var(H) - 2E_{var}Cov(H,S) + E^2_{var}Var(S)\Big) + O(n^{-2}). \tag{4.10}$$

(One may readily obtain a biased estimate of the variance from the sample [35]). Upon substitution of (4.1) into equations like the following:

$$Var(H) = \langle (\psi^*_t H \psi_t)^2 \rangle - \langle \psi^*_t H \psi_t \rangle^2 \quad , \tag{4.11}$$

$$Cov(H,S) = \langle (\psi^*_t H \psi_t)(\psi^*_t \psi_t) \rangle - \langle \psi^*_t H \psi_t \rangle \langle \psi^*_t \psi_t \rangle \quad , \tag{4.12}$$

one obtains:

$$Var(\hat{E}_{MC}) = n^{-1}\Big(\sum_{k,\ell} \lambda_k \lambda_\ell (E_k - E_o)(E_\ell - E_o)\langle \phi^2_o \phi_k \phi_\ell \rangle + O(\lambda^3)\Big) + O(n^{-2}). \tag{4.13}$$

Hence the leading terms in the expected value (4.9) and the variance (4.13) of $\hat{E}_{MC}$ (4.8) are second order in λ.

Details on the derivations of (4.7) and (4.13) appear in Appendix B.

In practice there are one or more adjustable parameters in ψ_t, say β, which should be optimized. This has been done by either minimizing $\tilde{E}_{MC}$ for a reasonably large value of n, as for example [6], or by minimizing the variance estimate, $\tilde{Var}(\hat{E}_{MC})$, as for example [7]. The former is rigorously justified by the variational theorem only for an infinite sample, while the latter has a statistical justification: the confidence interval for $\hat{E}_{var}$ is a minimum for the given sample.

Assume an infinite sample, so that the sample estimate of E_{var} and the variance are exact. By differentiating (4.7) with respect to β, the optimum parameter values, $\beta*$, obey the following equations:

$$\sum_k \lambda_k \frac{\partial \lambda_k}{\partial \beta_m}\bigg|_{\beta^*} (E_k - E_o) = 0, \text{ all m.} \tag{4.14}$$

By differentiating (4.13), the true variance of $\hat{E}_{MC}$,

$$\sum_{k,\ell} \lambda_k \frac{\partial \lambda_\ell}{\partial \beta_m}\bigg|_{\beta^{**}} (E_k - E_o)(E_\ell - E_o)\langle \phi_o^2 \phi_k \phi_\ell \rangle = 0, \text{ all m.} \tag{4.15}$$

Barring the trivial case of zero λ_k (for all k) and/or $\partial \lambda_k / \partial \beta_m$ (for all k and m), it is apparent that $\beta^* \neq \beta^{**}$. Hence, even if it were possible to optimize β by minimizing the true variance of $\hat{E}_{MC}$, the Virial Theorem would not be obeyed; additional scaling is required [7].

It is important to contrast (4.7) and (4.13), above, with their analogs for a normalized variational wavefunction

$$\psi_t = (1 + \sum_k \lambda_k{}^2)^{-\frac{1}{2}}(\phi_o + \sum_k \lambda_k \phi_k) \quad . \tag{4.16}$$

In this case the (crude) VMC estimate of E_{var} is given by

$$\hat{E}_{MC} = n^{-1}\sum_i H(R^{(i)}) \quad . \tag{4.17}$$

Now, the expected value of $\hat{E}_{MC}$ is the expected value of H which is E_{var}:

$$\langle\hat{E}_{MC}\rangle = E_{var} \quad . \tag{4.18}$$

Equation (4.17) is an unbiased estimate of E_{var}, whereas the ratio estimate (4.8) is biased (4.9).

In VMC calculations, where n is large, the unbiased estimate is not preferred; the variance of the ratio estimator involves the correlation between two random variables (H (4.3) and S (4.4)), which reduces its value (a fact well-known to survey statisticians [36]). This fact is also clear by comparing (4.13) with (4.19), which is obtained from (4.11), (4.16) and (4.17):

$$Var(\hat{E}_{MC}) = n^{-1}\left[E_o^2(\langle\phi_o^4\rangle - 1) + 2\sum_k \lambda_k E_o(E_o + E_k)\langle\phi_o^3\phi_k\rangle + 0(\lambda^2)\right] \tag{4.19}$$

(See Appendix C for details.) The variance of the unbiased estimate (4.17) is zeroth order in λ while that of the ratio estimate (4.8) is second order.

V. Diffusion Monte Carlo (DMC)

A recent and promising development is DMC, variations of which have been applied by several workers [25-27] plus earlier work cited therein. DMC contrasts with VMC by having the potential to estimate the exact ground state energy. We will focus on the approach of Reynolds et al. [25].

Consider the time dependent Schroedinger equation in imaginary time and in atomic units

$$-\frac{\partial \Phi}{\partial t}(R,t) = (H - E_T)\Phi(R,t) \qquad , \qquad (5.1)$$

where E_T is a constant shift in the zero of energy. The solution to this equation is given by

$$\Phi(R,t) = \sum_k \delta_k(t)\phi_k(R) \qquad , \qquad (5.2)$$

where δ_k is a time-dependent constant and ϕ_k is the exact wavefunction for the k-th state; δ_k decays exponentially in time such that the ground state (k = o) dominates as $t \longrightarrow \infty$ [25]:

$$\delta_k = c_k \exp\left(-(E_k - E_T)t\right) \qquad . \qquad (5.3)$$

DMC is concerned with the time dependent distribution function

$$f(\underset{\sim}{R},t) \equiv \Phi(\underset{\sim}{R},t)\psi_T(\underset{\sim}{R}) \qquad , \tag{5.4}$$

where ψ_T is an importance sampling or a "guiding" function. It is assumed that the nodes of ϕ_o exactly correspond to those of ψ_T so that f is non-negative for all $\underset{\sim}{R}$. The time dependence of f is given by

$$-\frac{\partial f}{\partial t} = -\tfrac{1}{2}\nabla^2 f + (E_L(\underset{\sim}{R}) - E_T)f + \underset{\sim}{\nabla}\cdot(f\psi_T^{-1}\underset{\sim}{\nabla}\psi_T) \qquad , \tag{5.5}$$

where E_L is the local energy of the guiding function:

$$E_L(\underset{\sim}{R}) \equiv \psi_T^{-1}H\psi_T \qquad . \tag{5.6}$$

The DMC algorithm yields, for sufficiently large t, a distribution of electron coordinates which may be viewed as a random sample drawn from the distribution function f.

Consider the expected value of E_L for the distribution

$$\langle E_L\rangle_f = \int E_L(\underset{\sim}{R})f(\underset{\sim}{R},t)d\underset{\sim}{R} \left/ \int f(\underset{\sim}{R},t)d\underset{\sim}{R} \right. \qquad . \tag{5.7}$$

Substitution of (5.2), (5.4) and (5.6) into (5.7) yields

$$\langle E_L\rangle_f = E_o + \sum_k \lambda_k \delta_k(t)\delta_o^{-1}(t)(E_k - E_o) + 0(\lambda^2\delta^2\delta_o^{-2}) \qquad , \tag{5.8}$$

where ψ_T has been expanded in terms of λ (4.1). Equation (5.8) shows that the exact energy is obtained in the limit of infinite time (D4).

Similarly, the expected value of $E_L{}^2$, that is,

$$\langle E_L^2 \rangle_f = \int E_L^2(\underline{R}) f(\underline{R},t) d\underline{R} \left/ \int f(\underline{R},t) d\underline{R} \right. \qquad , \qquad (5.9)$$

has the following leading terms:

$$\langle E_L^2 \rangle_f = E_o^2 + 2\sum_k \lambda_k \delta_k(t) \delta_o^{-1}(t) E_o(E_k - E_o) + \sum_k \lambda_k^2 (E_k - E_o)^2 + \ldots \qquad . \qquad (5.10)$$

Thus from (5.8) and (5.10)

$$Var(E_L)_f = \sum_k \lambda_k^2 (E_k - E_o)^2 + O(\lambda^2 \delta_m \delta_o^{-1}) + O(\lambda^3) \qquad . \qquad (5.11)$$

Details of the derivation of (5.8) and (5.10) and (5.11) appear in Appendix D.

DMC sampling involves generating M statistically independent blocks, each having a large number of configurations, $\underline{R}$. For our purposes we assume n independent configurations for each block.

The DMC energy estimate $\hat{E}_{DMC}$ is the average block energy, and each block energy is the average local energy for the n_c configurations in the block, where n_t time steps were required to reach the target time.

In obvious notation:

$$\hat{E}_{DMC} = M^{-1}\sum_{I} E^{(I)} \qquad , \qquad (5.12)$$

where

$$E^{(I)} = (n_c n_t)^{-1}\sum_{i} E_L\left(R^{(i,I)}\right) \qquad , \qquad (5.13)$$

where $R^{(i,I)}$ is the i-th configuration in the I-th block.

The expected value of $\hat{E}_{DMC}$ equals that of E_L (5.8):

$$\langle\hat{E}_{DMC}\rangle = E_o + \sum_{k=1}^{\infty} \lambda_k \delta_k(t) \delta_o^{-1}(t)(E_k - E_o) + \ldots \qquad ; \qquad (5.14)$$

The DMC energy estimate is unbiased only in the limit of infinite t, but the bias is negligible for large t.

From (5.10)-(5.13), the variance of the energy estimate is given by

$$Var(\hat{E}_{DMC}) = M^{-1}Var(E) \qquad\qquad (5.15a)$$

$$= M^{-1}(n_c n_t)^{-1}Var(E_L) \qquad\qquad (5.15b)$$

$$= M^{-1}(n_c n_t)^{-1}\left[\sum_{k=1}^{\infty} \lambda_k^2(E_k - E_o)^2 + \ldots\right] \qquad . \qquad (5.15c)$$

The leading term in the variance is second order in smallness parameters, independent of t. This provides a theoretical justification for the

empirical observation of Reynolds et al. [25] that the precision of DMC

calculations is strongly dependent on the choice of ψ_T.

VI. Applications--He Atom

VMC simulations were designed to compare results from minimizing the energy estimate $\hat{E}_{MC}$ (4.8) <u>versus</u> minimizing the sample estimate of the variance (4.10), $\hat{Var}(\hat{E}_{MC})$, with respect to the single variational parameter, β, in an uncorrelated wavefunction for 1^1S He:

$$\psi_t = \exp(-\beta s), \tag{6.1}$$

where Hylleraas coordinates are employed

$$s = r_1 + r_2, \quad t = r_1 - r_2, \quad u = r_{12} \quad . \tag{6.2}$$

In (6.2), r_i is the distance of the i-th electron from the nucleus, r_{ij} is the interelectronic distance, and

$$o \leqq t \leqq u \leqq s < \infty \quad . \tag{6.3}$$

After integrating over the Euler angles, the spatial element is given by

$$d\tau = 2\pi^2 u(s^2 - t^2)dsdtdu \quad , \tag{6.4}$$

and expressions for $\langle \psi_t H \psi_t \rangle$ and $\langle \psi_t \psi_t \rangle$ are given in [37].

Coordinates s, t, u were selected randomly using Importance sampling from a probability density function

$$\rho = (\alpha^3/\pi)\exp(-2\alpha s), \qquad (6.5)$$

with $\alpha = 2$. Importance sampling for Hylleraas coordinates [38] is described in Appendix E.

Since different random variables are used, the variance estimates for the crude and importance sampling samples will be different: the importance sampling energy estimate has a smaller variance. However the variance estimate of the ratio using importance sampling is still second order in λ (see Appendix B).

Subroutine VAO4A [39] is used to locate the minima in both the $\hat{E}_{MC}$ versus β and the $\hat{Var}(\hat{E}_{MC})$ versus β curves for 10 different random samples of 10^4 points. Table I contains values of $\hat{\beta}^*$ (estimate of β^* (4.14)) and $\hat{\beta}^{**}$ (estimate of β^{**} (4.15)) obtained from each sample and their averages. Also listed in Table I are the $\hat{E}_{MC}$ (E7) and the sample estimate (E10) of $Var(\hat{E}_{MC})$ (4.10) with corresponding $\hat{\beta}^*$ and $\hat{\beta}^{**}$ along with their standard deviations. Estimated errors are given for $\hat{Var}(\hat{E}_{MC})$; the formula used to calculate these is shown in Appendix F.

In all cases, $\hat{\beta}^*$ and $\hat{\beta}^{**}$, as calculated for the sample, do not coincide: the $\bar{\beta}^*$ and $\bar{\beta}^{**}$ do not overlap within one standard deviation. This provides numerical evidence to support the theoretical derivations (4.14) and (4.15) which show that $\beta^* \neq \beta^{**}$. Given that the precision of $\bar{\beta}^*$ is greater than that of $\bar{\beta}^{**}$ and that the energy estimates are

TABLE I Optimization of a One-Parameter Variational Function for 1 ^{1}S He[a].

Sample[b]	Minimize $\hat{E}_{MC}$			Minimize $\hat{\text{Var}}(\hat{E}_{MC})$			
	$\hat{\beta}$*[c]	$\hat{E}_{MC}$[d]	$\hat{\text{Var}}(\hat{E}_{MC})$[e]	$\hat{\beta}$**[f]	$\hat{E}_{MC}$[d]	$\hat{\text{Var}}(\hat{E}_{MC})$[e]	$\hat{\text{Var}}(\hat{\text{Var}}(\hat{E}_{MC}))$[g]
1	1.6875	-2.83 ± .04	1.8 ± .1 (-3)	1.6262	-2.83 ± .04	1.8 ± .2 (-3)	2.22 (-8)
2	1.7517	-2.93 ± .05	2.1 ± .2 (-3)	1.7164	-2.93 ± .05	2.1 ± .2 (-3)	4.18 (-8)
3	1.7210	-2.89 ± .05	2.1 ± .3 (-3)	1.6533	-2.89 ± .05	2.0 ± .2 (-3)	4.97 (-8)
4	1.7047	-2.85 ± .05	2.1 ± .4 (-3)	1.5291	-2.83 ± .04	1.8 ± .2 (-3)	3.65 (-8)
5	1.7171	-2.88 ± .05	2.1 ± .4 (-3)	1.5407	-2.86 ± .04	1.8 ± .2 (-3)	4.16 (-8)
6	1.7278	-2.90 ± .05	2.2 ± .5 (-3)	1.5456	-2.87 ± .04	1.9 ± .2 (-3)	4.45 (-8)
7	1.7329	-2.88 ± .05	2.2 ± .5 (-3)	1.5395	-2.85 ± .04	1.9 ± .2 (-3)	4.16 (-8)
8	1.6581	-2.84 ± .04	1.8 ± .2 (-3)	1.6123	-2.84 ± .04	1.9 ± .2 (-3)	3.42 (-8)
9	1.7334	-2.86 ± .05	2.0 ± .3 (-3)	1.6824	-2.84 ± .04	1.8 ± .2 (-3)	4.84 (-8)
10	1.7055	-2.86 ± .05	2.1 ± .4 (-3)	1.5304	-2.84 ± .04	1.8 ± .2 (-3)	3.72 (-8)

Averages Computed From the Above Samples:

$\overline{\hat{\beta}*}$ ⌐ 1.71 ± 0.03 $\overline{\hat{\beta}**}$ = 1.60 ± 0.07

$\overline{\hat{E}_{MC}}$ =-2.87 ± 0.03 $\overline{\hat{E}_{MC}}$ =-2.86 ± 0.03

$\overline{\hat{\text{Var}}(\hat{E}_{MC})}$ = 2.1 ± .1 (-3) $\overline{\hat{\text{Var}}(\hat{E}_{MC})}$ = 1.9 ± .2 (-3)

[a] Trial Wavefuction ψ_t is given by (6.1); Sampling density function given by (6.5) with α=2. True values are: β_{opt}=1.6875 and E_{var} = -2.848 au.

[b] Different samples of 10^4 random configuration points.

[c] $\hat{\beta}*$ estimates $\beta*$ defined by (4.14).

[d] $\hat{E}_{MC}$ defined by (4.8) in au with 1 standard deviation; see (E7).

[e] $\hat{\text{Var}}(\hat{E}_{MC})$ estimates (4.10) from sample variances and covariances with 1 standard deviation; see (E10).

[f] $\hat{\beta}**$ estimates $\beta**$ defined by (4.15).

[g] $\hat{\text{Var}}(\hat{\text{Var}}(\hat{E}_{MC}))$ defined by (F1), with appropriate use of $\tilde{H}$ and $\tilde{S}$ (E8).

comparable (both $\bar{\hat{E}}_{MC}$ values are within one standard deviation of E_{var}) there would seem to be no reason to choose the $Var(\hat{E}_{MC})$ as a minimization criterion.

Further studies using a 3-parameter Hylleraas wavefunction [37] also led to the same conclusions as those provided by the above example. While it was possible to use minimization of $\hat{E}_{MC}$ to optimize the wavefunction, it was found that the sensitivity of the $\hat{Var}(\hat{E}_{MC})$ to changes in the parameters made it expensive and impractical to use this optimization scheme. It was apparent in this case also that $\hat{\beta}^{**}$ would not be the same as $\hat{\beta}^{*}$ for a given sample.

An attempt was also made to optimize the normalized version of (6.1) using the $\hat{Var}(\hat{E}_{MC})$ criterion. This proved to be impossible, as $\hat{\beta}^{**}$ consistently went to very large values, making ψ_t zero and hence $\hat{Var}(\hat{E}_{MC})$ zero also. As this result is due, in part, to the extreme simplicity of the wavefunction (6.1), it is difficult to extrapolate to the case of a more complicated, normalized ψ_t. However, these results concur with the theoretical result (4.19), which being zero-th order in λ, would provide a different estimate of β^{**} than would (4.13). It is also of note that (4.19) predicts a larger variance than (4.13) which is in practice true, since the ratio estimator uses correlation between the numerator and denominator to reduce its variance.

VII. van der Waals Interaction Energy: An Electron Correlation Problem

(a) Introduction [9, 40-46]

van der Waals (vdW) molecules are weakly bound complexes of atoms or molecules which exhibit small dissociation energies (arbitrarily <0.09 hartree (40 kJ mol^{-1})) and large bond lengths (0.2-0.5 x 10^{-9} m). They are held together not by chemical binding forces but by predominantly physical, multipolar forces. van der Waals molecules may be found in gases at low temperatures and in electron-beam plasmas; they modify virial and transport properties of gases, chemical reaction mechanisms, nucleation phenomena and absorption in rare-gas plasmas.

Since so few of these molecules are present in a given system, and they are relatively unstable, conventional experiments that study dissociation energies, bond lengths etc., are not easily applied: spectra have diffuse bands, small responses must be separated out, and so on. There remains a need for theoretical calculations to supplement experimental results and to make predictions where experiments cannot afford information.

Of particular interest is the van der Waals interaction energy, ΔE_{INT}. Given an ab initio quantum mechanical calculation, there are two approaches to finding ΔE_{INT}: Variational calculations calculate ΔE_{INT} as a difference while Perturbation theory calculates ΔE_{INT} directly as a sum of components.

Considering a vdW dimer, the variational scheme is represented below--ΔE_{INT} is the difference between the total energy of the

"supermolecule" E_{AB} (consisting of the interacting subsystems A and B

separated by distance R) and the sum of the non-interacting subsystem

energies E_A and E_B.

$$\Delta E_{INT}(R) = E_{AB}(R) - \left(E_A(R = \infty) + E_B(R = \infty) \right) \quad . \qquad (7.1)$$

When R is the equilibrium distance R_e, $\Delta E_{INT}(R_e)$ is the van der Waals

binding energy. In general, if neutral, non-polar molecules/atoms A and

B are considered, the existence of a minimum in the inter-molecular/atomic

energy versus distance R curve is almost entirely a manifestation of

inter-molecular/atomic correlation energy. (In the Perturbation scheme,

this is called the dispersion energy.) Correlation energy arises

because coulombic repulsion between like charges tends to keep electrons

apart, thus lowering the total energy of a system. In the simple

Hartree-Fock Self-Consistent-Field (HF-SCF) theory, Fermi correlation (a

consequence of the Pauli Exclusion Principle--two electrons with like spin

cannot exist in the same spatial orbital) is accounted for by requiring

the molecular wavefunction to be antisymmetric. However, correlation

between electrons with unlike spins (Coulombic correlation) is not

considered and so the term correlation energy refers to the difference

between the true energy and the Hartree-Fock energy E_{HF}.

Typically, correlation energy makes up less than one percent of

the total energy and although for qualitative purposes it can often be

ignored, properties of chemical interest require its recovery for quantitative results. One major problem encountered in the variational treatment is the Basis Set Superposition Error (BSSE) [47-49] which results in a ΔE_{INT} that is too large. This is a purely mathematical artifact: E_{AB} is calculated with a qualitatively better basis set than the ones from which E_A and E_B are found, simply because basis set AB is a union of basis sets A and B. Instead of the costly procedure of expanding the subsystem basis sets to the point where the effect of doubling the size at R_e does not radically improve E_{AB}, counterpoise correction or ghost orbitals [48, 49] may be used. Here, E_A and E_B are computed with basis set AB; that is, the E_A computation is allowed use of basis set B, but without B actually being present. Studies have shown that such techniques do not overcorrect [49].

Configuration Interaction (CI) calculations mix contributions from excited states of the same symmetry with the ground state in order to recover the correlation energy. The use of explicitly correlated wavefunctions introduces interelectronic distances r_{ij} into the calculation so that when r_{ij} is small, the wavefunction ψ is small.

To get accurate absolute values of the total energy of a quantum mechanical system is difficult; the absolute error in, say, the He_2 problem is greater than the well-depth $\Delta E_{INT}(R_e)$. The Variational supermolecule approach hinges on the assumption that the error in the variationally-obtained energy remains constant as a function of the internuclear separation so that cancellation of errors occurs upon subtraction of the quantities in (7.1). In the He_2 problem, the ground

state energy is on the order of -5.8 hartree while the well-depth is about $-3.2-3.8 \times 10^{-5}$ hartree, and so the absolute error in the variationally-obtained energies must remain constant to about 1 ppm in order to get accurate values of ΔE_{INT}.

The He_2 problem was the first vdW molecule for which good results were obtained in the early 1970's using CI or MCSCF calculations. Schaefer and McLaughlin [50] made the assumption that intra-atomic correlation remains constant over all separations R. Bertoncini and Wahl [51] included small changes in the intra-atomic correlation energy; Liu and McLean [52] also incorporated coupling between inter- and intra-atomic components. Dacre [53] attempted to remove the Basis Set Superposition Error in addition to incorporating inter- and intra-atomic correlation. Burton [54] used CEPA-PNO correlated wavefunctions.

Coldwell and Lowther [7] used a very complicated explicitly correlated wavefunction in a Variational Quantum Monte Carlo calculation. Their result for ΔE_{INT} at R_e = 5.6 au was -3.55 (± 0.15) $\times 10^{-5}$ hartree as compared to Burton's -3.339×10^{-5} hartree at R_e = 5.63 au and to the experimental differential-scattering cross-section measurement result of -3.35×10^{-5} hartree at R_e = 5.6 au of Burgmans, Farrar and Lee [55]. In order to obtain such quantitative results for a VMC calculation, sample sizes of 377,000 and 782,000 points were used.

In the second part of this work, simple explicitly correlated wavefunctions are examined in a Variational Monte Carlo calculation in order to see whether a qualitative estimate of ΔE_{INT} for He_2 can be made. Quantitative results are not expected as the wavefunctions employed here

have only one variational parameter as compared to the hundreds of parameters used in Coldwell and Lowther's calculation.

Given a partitioning of the Hamiltonian which eliminates the need for evaluating gradients of an SCF determinant, and the use of well-known HF-SCF orbital energies, formulae are derived which allow the method to be applied to He_2. A method of Importance sampling in prolate spheroidal coordinates is derived and the whole incorporated into a computer program.

Results from the first part of this work are used to optimize the correlated wavefunctions.

VIII. Theory

(a) Partitioning of the Hamiltonian-Combining Results from HF-SCF and Correlation Functions

From HF-SCF theory [56], the i-th spin orbital ϕ_i is an eigenfunction of the effective Hamiltonian H_i^{eff}:

$$H_i^{eff} \phi_i(1) = \varepsilon_i \phi_i(1) \qquad , \qquad i = 1, \ldots, N \qquad (8.1)$$

where ϕ_i is a product of a spatial function f_i and spin component α or β:

$$\phi_i(x) = f_i(r) \begin{cases} \alpha(s) \\ \beta(s) \end{cases} \qquad (8.2)$$

and 1 represents the space and spin coordinates of electron 1, ε_i is the orbital energy. There are N such equations for the N orthonormal spin-orbitals in the N-electron system. Thus the expectation value of H_i^{eff} is ε_i:

$$\langle \phi_i(1) | H_i^{eff} | \phi_i(1) \rangle = \varepsilon_i \qquad , \qquad i = 1, \ldots, N \qquad . \qquad (8.3)$$

More explicitly,

$$\langle \phi_i(1) | H_i^{eff} | \phi_i(1) \rangle \equiv \langle \phi_i(1) | H_i | \phi_i(1) \rangle + J_i - K_i,$$

$$i = 1, \ldots, N \qquad (8.4)$$

where

$$H_i(1) = -\tfrac{1}{2}\nabla_1^2 - \sum_\alpha Z_\alpha/r_{1\alpha} \qquad (8.5)$$

contains the electronic kinetic energy and electron-nuclear attraction operators (atomic units are $\hbar = m_e = e = 1$). The J_i and K_i are the coulomb and exchange integrals, respectively:

$$J_i = \sum_{j\neq i} J_{ij} = \sum_{j\neq i} \iint \phi_i^*(1)\phi_j^*(2)r_{12}^{-1}\phi_i(1)\phi_j(2)dx_1 dx_2 \qquad (8.6)$$

and

$$K_i = \sum_{j\neq i} K_{ij} = \sum_{j\neq i} \iint \phi_i^*(1)\phi_j^*(2)r_{12}^{-1}\phi_j(1)\phi_i(2)dx_1 dx_2 \quad . \qquad (8.7)$$

Define coulomb and exchange operators J_i and K_i such that,

$$J_i(1)\phi_i(1) = \left[\sum_{j\neq i}\int \phi_j^*(x)r_{1x}^{-1}\phi_j(x)dx\right]\phi_i(1) \qquad (8.8)$$

and

$$K_i(1)\phi_i(1) = \left[\sum_{j\neq i}\int \phi_j^*(x)r_{1x}^{-1}P_{ij}\phi_j(x)dx\right]\phi_i(1)$$

$$= \left[\sum_{j\neq i}\int \phi_j^*(x)r_{1x}^{-1}\phi_i(x)dx\right]\phi_j(1) \qquad (8.9)$$

where P_{ij} permutes orbitals i and j. Equations (8.8) and (8.9) can be related to integrals J_i (8.6) and K_i (8.7) respectively by

$$J_i = \langle \phi_i^*(1) | J_i(1) | \phi_i(1) \rangle \tag{8.10}$$

and

$$K_i = \langle \phi_i^*(1) | K_i(1) | \phi_i(1) \rangle \quad . \tag{8.11}$$

Hence,

$$H_i^{eff} = H_i + J_i - K_i \quad . \tag{8.12}$$

Summing H_i^{eff} over all i = 1, ..., N gives the total Hartree-Fock Hamiltonian, H_{HF}^{TOT},

$$H_{HF}^{TOT} = \sum_{i=1}^{N} H_i + \sum_{i=1}^{N} (J_i - K_i) \quad . \tag{8.13}$$

The terms in (8.13) accounting for interelectronic interactions make up the Hartree-Fock potential V_{HF} [57],

$$V_{HF} = \sum_{i=1}^{N} (J_i - K_i) \quad . \tag{8.14}$$

A Slater determinant $\Phi(\underline{x})$ is the simplest, totally antisymmetric, N-particle wavefunction and is a (normalized) determinant of N spin orbitals

$$\Phi(\underset{\sim}{x}) = (N!)^{-\frac{1}{2}} \sum_{r=1}^{N!} (-1)^r P_r \big(\phi_1(1)\phi_2(2) \ldots \phi_N(N)\big) \qquad . \qquad (8.15)$$

Equation (8.15) can also be written in terms of the antisymmetrization operator, A:

$$A = (N!)^{-\frac{1}{2}} \sum_{r=1}^{N!} (-1)^r P_r \qquad\qquad (8.16)$$

and the Hartree product ϕ_{HP}

$$\phi_{HP} = \phi_1(1) \ldots \phi_N(N) \qquad\qquad (8.17)$$

so that

$$\Phi(\underset{\sim}{x}) = A\phi_{HP} \qquad . \qquad\qquad (8.18)$$

Following Appendix C of Blinder [56], it can be shown that

$$\langle \Phi(\underset{\sim}{x}) | H_{HF}^{TOT} | \Phi(\underset{\sim}{x}) \rangle = \langle \phi_{HP} | H_{HF}^{TOT} | \sum_{r=1}^{N!} (-1)^r P_r \phi_{HP} \rangle$$

$$= \sum_{i=1}^{N} \langle \phi_i | H_i | \phi_i \rangle + \sum_{i=1}^{N} (J_i - K_i) = \sum_{i=1}^{N} \varepsilon_i \qquad . \qquad (8.19)$$

(see Appendix G for details of (8.19)).

In (8.19), interactions between electron pairs have been counted twice. Rearranging (8.19) so as to include only unique interactions,

$$\sum_{i=1}^{N} <\phi_i|H_i|\phi_i> + \sum_{j>i} \sum_{i=1}^{N} (J_{ij} - K_{ij}) = \sum_{i=1}^{N} \varepsilon_i - \sum_{j>i} \sum_{i=1}^{N} (J_{ij} - K_{ij}) \quad .$$

$$(8.20)$$

It is a well-known result [56] that the right-hand side of (8.20) is the Hartree-Fock energy, E_{HF}

$$E_{HF} = \sum_{i=1}^{N} \varepsilon_i - \sum_{j>i} \sum_{i=1}^{N} (J_{ij} - K_{ij}) \quad . \qquad (8.21)$$

Since from the definition of V_{HF} (8.14) and the result in (8.19),

$$<\Phi(\underset{\sim}{x})|V_{HF}|\Phi(\underset{\sim}{x})> = <V_{HF}> = \sum_{j\neq i} \sum_{i=1}^{N} (J_{ij} - K_{ij}) \qquad (8.22)$$

the Hartree-Fock energy can be expressed from (8.21) and (8.22) as

$$E_{HF} = \sum_{i=1}^{N} \varepsilon_i - \tfrac{1}{2}<V_{HF}> \quad . \qquad (8.23)$$

Gibbs and Dunn [6] have used the definition of the V_{HF} (8.14) in order to partition the electronic Hamiltonian in a convenient fashion for use with explicitly correlated wavefunctions. Consider the electronic Hamiltonian H

$$H = -\tfrac{1}{2}\sum_{i}\nabla_i^2 - \sum_{i}\sum_{\alpha}Z_{\alpha}r_{i\alpha}^{-1} + \sum_{j>i}\sum_{i}r_{ij}^{-1}$$

where the first term represents the kinetic operator, the second the electron-nuclear attraction and the third, the electron-electron repulsion. Indices i and j refer to summations over electrons; index α refers to summation over the nuclei. The Hamiltonian can be partitioned as follows:

$$H = \left[H - \sum_{j>i} \sum_i r_{ij}^{-1} + V_{HF} \right] + \left[\sum_{j>i} \sum_i r_{ij}^{-1} - V_{HF} \right] \tag{8.24}$$

$$= H_{HF}^{TOT} + H_1$$

(from (8.13) and (8.14)) since addition and subtraction of terms leaves H unchanged.

Employing $\Phi(\underset{\sim}{x})$ to evaluate $\langle H \rangle$ will give the Hartree-Fock energy for the system. Knowing this, and using (8.24)

$$E_{HF} = \langle \Phi(\underset{\sim}{x}) | H | \Phi(\underset{\sim}{x}) \rangle = \langle H_{HF}^{TOT} + H_1 \rangle = \sum_{i=1}^{N} \varepsilon_i + \langle H_1 \rangle \quad . \tag{8.25}$$

Comparison of (8.25) and (8.23) shows that

$$\langle H_1 \rangle = \langle \sum_{j>i} \sum_i r_{ij}^{-1} - V_{HF} \rangle = -\tfrac{1}{2} \langle V_{HF} \rangle \tag{8.26}$$

and so for Hartree-Fock wavefunctions,

$$\langle \sum_{j>i} \sum_i r_{ij}^{-1} \rangle = \tfrac{1}{2} \langle V_{HF} \rangle \tag{8.27}$$

as mentioned by Gibbs and Dunn [6].

Consider now an explicitly correlated wavefunction of the form

$$\psi(\underline{x}) = U(\underline{r}_{ij}) \Phi(\underline{x}) \tag{8.28}$$

where as before $\Phi(\underline{x})$ is a Hartree-Fock wavefunction and $U(\underline{r}_{ij})$ is a correlation function in terms of the interelectronic distances $\underline{r}_{ij}$. The functional form of $U(\underline{r}_{ij})$ will be chosen to be

$$U(\underline{r}_{ij}) = \exp\{-u\} = \exp\left\{-\sum_{j>i}\sum_{i} u(r_{ij})\right\} \tag{8.29}$$

The variational energy using $\psi(\underline{x})$ (8.28) and the partitioned form of the Hamiltonian (8.24) is

$$E_{var} = \langle \psi^*(\underline{x}) | H_{HF}^{TOT} | \psi(\underline{x}) \rangle \Big/ \langle \psi^*(\underline{x}) \psi(\underline{x}) \rangle$$

$$+ \langle \psi^*(\underline{x}) | H_1 | \psi(\underline{x}) \rangle \Big/ \langle \psi^*(\underline{x}) \psi(\underline{x}) \rangle \qquad . \tag{8.30}$$

The numerator of the first term in (8.30) is

$$\langle H_{HF}^{TOT} \rangle = \langle (U\Phi)^* \Big| -\tfrac{1}{2}\sum_{i}\nabla_i^2 - \sum_{i}\sum_{\alpha} Z_\alpha r_{i\alpha}^{-1} + V_{HF} \Big| U\Phi \rangle \qquad . \tag{8.31}$$

Substituting (8.29) for U in (8.31) leads to

$$\langle H_{HF}^{TOT} \rangle = \sum_i \epsilon_i \langle \psi^* \psi \rangle - \frac{1}{2} \sum_i \langle (U\Phi)^* \Big| 2\vec{\nabla}_i e^{-u} \cdot \vec{\nabla}_i \Phi + \Phi \vec{\nabla}_i^2 e^{-u} \rangle \quad . \qquad (8.32)$$

It is shown in Appendix H that by applying the Divergence Theorem, the $\vec{\nabla}_i u \cdot \vec{\nabla}_i \Phi$ terms which arise from (8.32) disappear, leaving

$$\langle H_{HF}^{TOT} \rangle = \sum_i \epsilon_i \langle \psi^* \psi \rangle + \frac{1}{2} \sum_{i=1}^{N} \langle \vec{\nabla}_i u \cdot \vec{\nabla}_i u \rangle \quad . \qquad (8.33)$$

Substituting (8.33) into (8.30), the final result is

$$E_{var} = \sum_i \epsilon_i + \langle H_1 \rangle / \langle \psi^* \psi \rangle + \frac{1}{2} \sum_i \langle \vec{\nabla}_i u \cdot \vec{\nabla}_i u \rangle / \langle \psi^* \psi \rangle \qquad (8.34)$$

as in Gibbs and Dunn [6].

The advantages of (8.34) are clear: the HF-SCF orbital energies are well-known; $\langle H_1 \rangle$ involves only multiplicative terms; the gradient of the determinant $\Phi(\underline{x})$ need not be computed; $\vec{\nabla}_i u$ are easily evaluated.

(b) Evaluation of $\langle H_1 \rangle$

The calculation of $\langle H_1 \rangle$ in (8.34) involves finding the expectation value over a distribution function determined by the correlated wavefunction $\psi(\underline{x})$ (8.28). Following the approach in Appendix G, employing the properties of the antisymmetrizer, the Hermitian property of quantum mechanical operators and the symmetry of

$$\sum_{j>i} \sum_{i} r_{ij}^{-1}, \; V_{HF} \text{ and } U(\underline{r}_{ij})$$

in the coordinates, the unnormalized $\langle H_1 \rangle$ is written

$$\langle H_1 \rangle = \langle \phi_{HP} U^2(\underline{r}_{ij}) \Big| \sum_{j>i} \sum_{i} r_{ij}^{-1} - \sum_{i}(J_i - K_i) \Big| \sum_{r=1}^{N!} (-1)^r P_r \phi_{HP} \rangle$$

$$= \langle \sum_{j>i} \sum_{i} r_{ij}^{-1} - V_{HF} \rangle \qquad . \tag{8.35}$$

The presence of $U(\underline{r}_{ij})$, while explicitly introducing interelectronic coordinates which effectively account for Coulombic correlation missing from the HF-SCF scheme, complicates the evaluation of molecular integrals. It is no longer possible to separate multiple integrals; it is also true that while spin-orthogonality allows for the elimination of many integrals, those that contain spatially orthogonal orbitals also contain a factor of U^2 and so are non-zero. While expressions such as (8.35) may not be tractable analytically, Monte Carlo Methods can quite easily be used to numerically compute many-dimensional, inseparable integrals.

It is always desirable from the point of efficiency to identify zero integrals and to eliminate them prior to computation. Many of the permutations of ϕ_{HP} in (8.35) cause zero integrals due to spin-orthogonality. The expansions from the Slater determinant rapidly increase in size with N, many terms lead to superfluous integrals. It is possible to employ a product of determinants containing orbitals of like spins [25,26] to bypass the construction of such terms.

Instead of (8.28) use

$$\psi(\underline{x}) = U(\underline{r}_{ij}) d_\alpha \left[(r,\alpha)_k \right] d_\beta \left[(r,\beta)_k \right] \qquad . \qquad (8.36)$$

where d_s represents a determinant of orbitals with spin s

$$d_s = \det \left[\phi_e \left[(r,s)_k \right] \right] , \qquad k, e \leq N/2 \quad . \qquad (8.37)$$

For example, consider the two-electron problem: the Slater determinant gives

$$\Phi(\underline{r},s) \propto \phi_1(r_1,\alpha)\phi_2(r_2,\beta) - \phi_2(r_1,\beta)\phi_1(r_2,\alpha) \qquad . \qquad (8.38)$$

All integrals involving $\phi_2(1)\phi_1(2)$ disappear due to spin orthogonality, that is,

$$\langle \phi_1(1)\phi_2(2)U^2(r_{12}) | \mathcal{O} | \phi_2(1)\phi_1(2) \rangle = 0 \quad . \qquad (8.39)$$

The use of (8.36) uses only

$$d_\alpha\big((r,\alpha)_1\big)d_\beta\big((r,\beta)_2\big) = \phi_1(r_1,\alpha)\phi_2(r_2,\beta) \quad . \tag{8.40}$$

For a four-electron, closed-shell system the Hartree product is

$$\phi_{HP}(x_1,x_2,x_3,x_4) = \phi_1(r_1,\alpha)\phi_2(r_2,\beta)\phi_3(r_3,\alpha)\phi_4(r_4,\beta) \tag{8.41}$$

and (8.36) gives

$$\psi(\underset{\sim}{x}) = U(\underset{\sim}{r}_{ij}) \begin{vmatrix} \phi_1(1)\phi_1(3) \\ \phi_3(1)\phi_3(3) \end{vmatrix} \times \begin{vmatrix} \phi_2(2)\phi_2(4) \\ \phi_4(2)\phi_4(4) \end{vmatrix}$$

$$= U(\underset{\sim}{r}_{ij})\det(\phi_1\phi_3)\det(\phi_2\phi_4) \quad . \tag{8.42}$$

The expansion of (8.42) leads to four terms involving ϕ_i:

$$\psi(\underset{\sim}{x}) = U(\underset{\sim}{r}_{ij})[\phi_1(1)\phi_2(2)\phi_3(3)\phi_4(4) - \phi_1(1)\phi_4(2)\phi_3(3)\phi_2(4)$$

$$- \phi_3(1)\phi_2(2)\phi_1(3)\phi_4(4) + \phi_3(1)\phi_4(2)\phi_1(3)\phi_2(4)] \tag{8.43}$$

which arise from permutations between orbitals with like spin. The full 4×4 Slater determinant produces $4! = 24$ terms; the 20 additional terms to those in (8.43) contain permutations among orbitals with unlike spin. After operation of an operator which does not affect spin, multiplication by ϕ_{HP}, these terms will result in zero integrals.

Then from (8.35) and (8.43), for the four-electron system,

$$\langle H_1 \rangle_{4e^-} = \langle \phi_{HP} U^2(r_{ij}) \left| \sum_{j>i} \sum_{i=1}^{4} r_{ij}^{-1} - \sum_{i=1}^{4} (J_i - K_i) \right|$$

$$\times \quad \phi_1(1)\phi_2(2)\phi_3(3)\phi_4(4) - \phi_1(1)\phi_4(2)\phi_3(3)\phi_2(4)$$

$$- \phi_3(1)\phi_2(2)\phi_1(3)\phi_4(4) + \phi_3(1)\phi_4(2)\phi_1(3)\phi_2(4) \rangle . \quad (8.44)$$

IX. Monte Carlo Evaluation of E_{var} for $^1\Sigma_g^+ He_2$

(a) Coordinate System [58]

Figure I shows the prolate spheroidal system of coordinates for He_2:

$$\xi = R^{-1}(r_A + r_B), \qquad 1 \leq \xi < \infty \tag{9.1a}$$

$$\eta = R^{-1}(r_A - r_B), \qquad -1 \leq \eta \leq 1 \tag{9.1b}$$

$$0 \leq \phi < 2\pi \tag{9.1c}$$

where r_N is the distance of the electron from nucleus N and R is the internuclear separation. The interelectronic distance r_{ij} is taken from

$$r_{ij}^2 = (R/2)^2 \left[\xi_i^2 + \eta_i^2 + \xi_j^2 + \eta_j^2 - 2\xi_i\eta_i\xi_j\eta_j - 2 \right.$$

$$\left. -2\left[(\xi_i^2 - 1)(1 - \eta_i^2)(\xi_j^2 - 1)(1 - \eta_j^2)\right]^{\frac{1}{2}} \cos(\phi_i - \phi_j) \right] \tag{9.2}$$

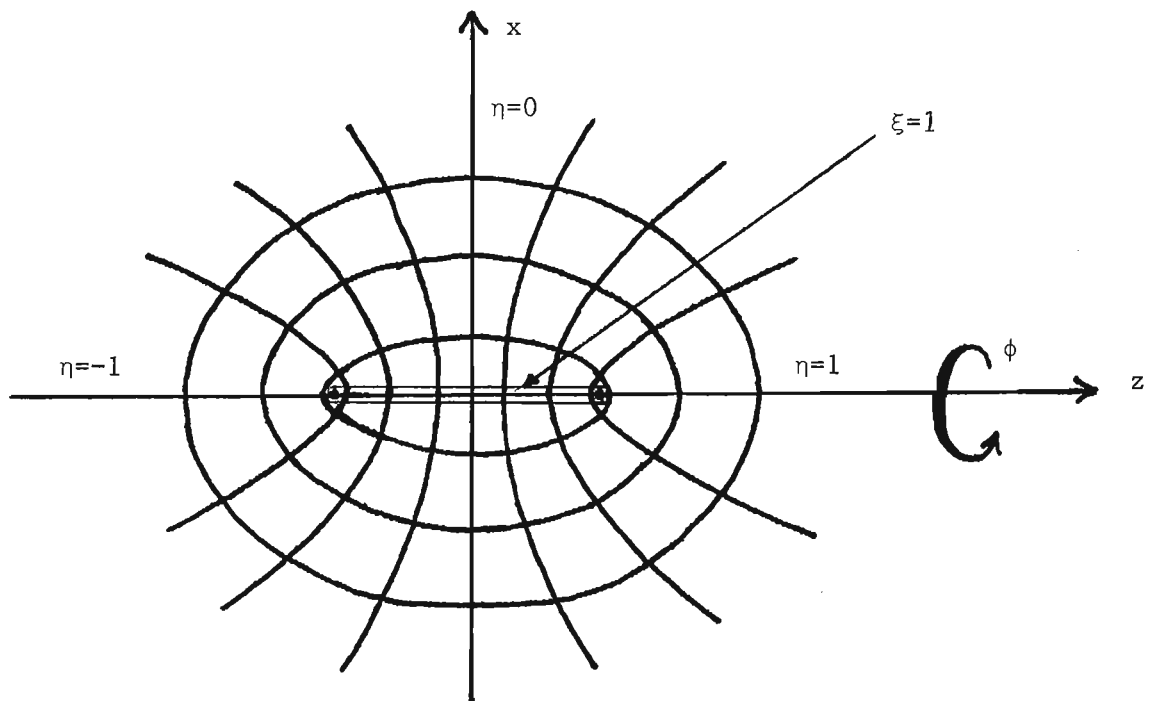The volume element is $(R^3/8)(\xi^2 - \eta^2)d\xi d\eta d\phi$.
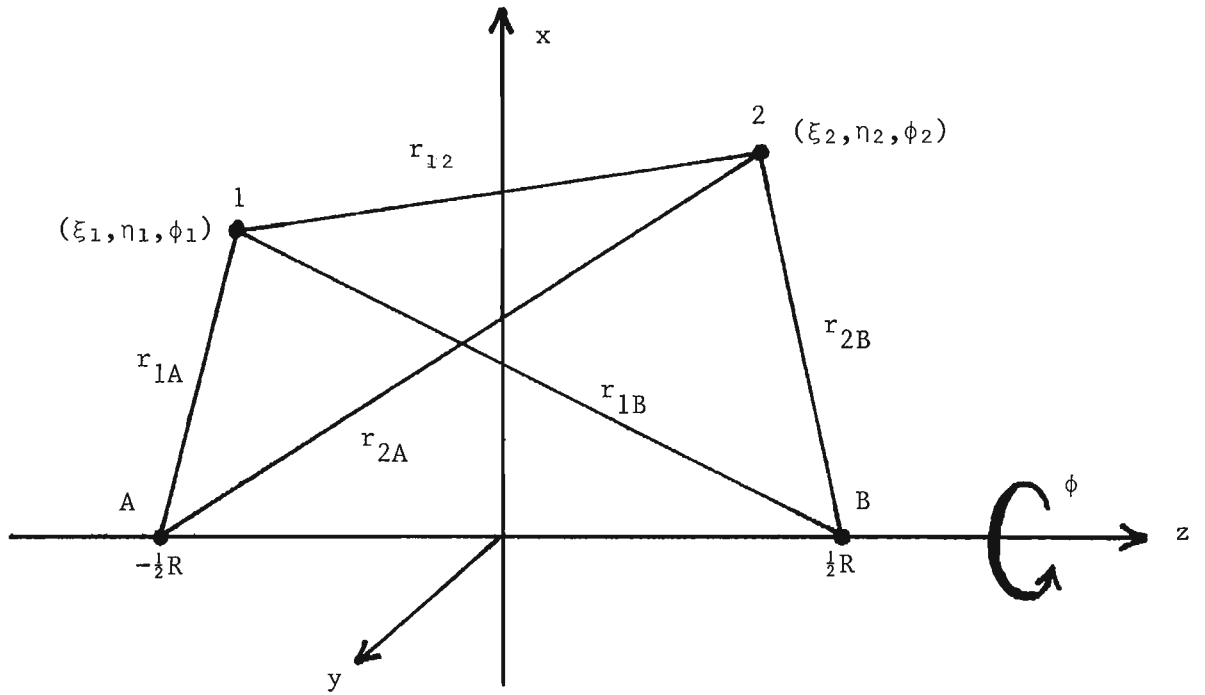
The following give the transformation into Cartesian coordinates:

$$x = \tfrac{1}{2}R\left[(\xi^2 - 1)(1 - \eta^2)\right]^{\frac{1}{2}}\cos\phi \qquad\qquad (9.3a)$$

$$y = \tfrac{1}{2}R\left[(\xi^2 - 1)(1 - \eta^2)\right]^{\frac{1}{2}}\sin\phi \qquad\qquad (9.3b)$$

$$z = \tfrac{1}{2}R\xi\eta \qquad\qquad (9.3c)$$

Figure I. Prolate Spheroidal Coordinate System

(b) Basis Set and Molecular Orbitals--Evaluation of $V_{HF}|\det(\phi_1\phi_3)\det(\phi_2\phi_4)\rangle$

For the ground-state He_2 vdW molecule, the minimal basis set of bonding and anti-bonding molecular spin-orbitals, ϕ_i, $i = 1,...,4$, can be expressed as linear combinations of the normalized 1s atomic orbitals χ_A and χ_B centered on nuclei A and B respectively, where

$$\chi_N = (\zeta^3/\pi)^{\frac{1}{2}}\exp(-\zeta r_N) \quad . \tag{9.4}$$

The ζ is the orbital exponent.

At internuclear separations where χ_A and χ_B contribute equally to ϕ_i, let

$$\phi_1 = c_1(\chi_A + \chi_B)\alpha \tag{9.5a}$$

$$\phi_2 = c_1(\chi_A + \chi_B)\beta \tag{9.5b}$$

$$\phi_3 = c_2(\chi_A - \chi_B)\alpha \tag{9.5c}$$

$$\phi_4 = c_2(\chi_A - \chi_B)\beta \tag{9.5d}$$

where ϕ_1, ϕ_2 represent bonding and ϕ_3, ϕ_4 anti-bonding spin-orbitals.

The orbital energies ε_i and coefficients c_1, c_2 are easily found from HF-SCF programs such as HONDO [59]. It is then straight-forward to calculate the $\langle\sum_{j>i}\sum_i r_{ij}^{-1}\rangle$ portion in $\langle H_1\rangle$ (8.44), $\sum_i\langle\vec{\nabla}_i u\cdot\vec{\nabla}_i u\rangle$ (8.34) and the

normalization factor $\langle \psi^* \psi \rangle$ within the Monte Carlo algorithm (IXc) given a method of sampling in configuration space and formulae for the operators, $U(\underline{r}_{ij})$ and ϕ_i. However, evaluation of

$$\langle \sum_{i=1}^{4} (J_i - K_i) \rangle$$

involves expressing J_i, K_i in terms of known analytic functions arising from atomic orbitals χ_A and χ_B.

Consider J_1; from (8.8)

$$J_1(1) = \int \phi_2^*(x) r_{1x}^{-1} \phi_2(x) dx + \int \phi_3^*(x) r_{1x}^{-1} \phi_3(x) dx$$

$$+ \int \phi_4^*(x) r_{1x}^{-1} \phi_4(x) dx \quad . \tag{9.6}$$

Substitution of (9.5b-d) into (9.6) and allowing that χ_A and χ_B are real, gives

$$J_1(1) = (c_1^2 + 2c_2^2) \left[\int \chi_A^2(r_x) r_{1x}^{-1} dr_x + \int \chi_B^2(r_x) r_{1x}^{-1} dr_x \right]$$

$$+ 2(c_1^2 - 2c_2^2) \int \chi_A(r_x) \chi_B(r_x) r_{1x}^{-1} dr_x \quad . \tag{9.7}$$

where the spin-components have been integrated over, and terms collected.

A like expression is found for $J_2(1)$ and

$$J_4(1) = J_3(1) = (2c_1{}^2 + c_2{}^2)\left[\int \chi_A{}^2(r_x)r_{1x}{}^{-1}dr_x + \int \chi_B{}^2(r_x)r_{1x}{}^{-1}dr_x\right]$$

$$+ 2(2c_1{}^2 - c_2{}^2)\int \chi_A(r_x)\chi_B(r_x)r_{1x}{}^{-1}dr_x \quad . \tag{9.8}$$

Roothan [60] gives a formula for

$$V_{1s} = \int \chi_{1s}{}^2(r_x)r_{1x}{}^{-1}dr_x \quad :$$

the integration over coordinate r_x amounts to finding the potential for electron 1 at its instantaneous position in space due to the charge distribution $\chi_{1s}{}^2$. This potential is, for a $(1s)^2$ charge distribution,

$$V_{1s}(1) = r_1{}^{-1}\left[1 - (1 + \zeta r_1)\exp(-2\zeta r_1)\right] \tag{9.9}$$

where r_1 is the electron-nuclear separation.

Rüdenberg [61] gives general formulae for two-center exchange integrals for Slater-type atomic orbitals which can be modified somewhat to calculate

$$\int \chi_A(r_x)\chi_B(r_x)r_{1x}{}^{-1}dr_x.$$

It is shown in Appendix I that, for $\chi_A\chi_B = 1s_A1s_B$,

$$\int \chi_A(r_x)\chi_B(r_x)r_{1x}^{-1}dr_x$$

$$= 2R^{-1}(R\zeta)^3 \sum_{\ell=0}^{\infty} \left[(2\ell + 1)/2\right]^{\frac{1}{2}} P_\ell(\eta_1) \sum_{n=0}^{2} \sum_{j=0}^{2} \omega_{nj} B_j^{0\ell}(0)$$

$$\times \left[Q_\ell(\xi_1)K_{\ell n}(\xi_1,\zeta R) + P_\ell(\xi_1)L_{\ell n}(\xi_1,\zeta R)\right] \qquad (9.10)$$

where P_ℓ and Q_ℓ are the Legendre polynomials of the first and second kind, respectively; $K_{\ell n}$ and $L_{\ell n}$ are given by:

$$K_{\ell n}(\xi_1,\alpha) = \int_{1}^{\xi_1} \exp(-\alpha\xi)\xi^n P_\ell(\xi)d\xi \qquad (9.11)$$

and

$$L_{\ell n}(\xi_1,\alpha) = \int_{\xi_1}^{\infty} \exp(-\alpha\xi)\xi^n Q_\ell(\xi)d\xi \qquad . \qquad (9.12)$$

Coefficients ω_{nj} are given in Table I of Rüdenberg [61] and $B_j^{0\ell}(0)$ are defined as

$$B_j^{0\ell}(0) = \left[(2\ell + 1)/2\right]^{\frac{1}{2}} \int_{-1}^{1} P_\ell(\eta)\eta^j d\eta \qquad . \qquad (9.13)$$

In a similar manner, $K_1\phi_i$ is from (8.9)

$$K_1(1)\phi_1(1) = \left[\int \phi_2^*(x)r_{1x}^{-1}\phi_1(x)dx\right]\phi_2(1) + \left[\int \phi_3^*(x)r_{1x}^{-1}\phi_1(x)dx\right]\phi_3(1)$$

$$+ \left[\int \phi_4^*(x)r_{1x}^{-1}\phi_1(x)dx\right]\phi_4(1) \qquad . \qquad (9.14)$$

Now, the first and third integrals in (9.14) are zero due to spin-orthogonality; only the term which permutes orbitals with like spin remains. Substituting (9.5b) – (9.5d) into (9.14) and integrating over spin leaves

$$K_1(1)\phi_1(1) = c_1 c_2 \left\{ \int \chi_A^2(r_x) r_{1x}^{-1} dr_x - \int \chi_B^2(r_x) r_{1x}^{-1} dr_x \right\} \phi_3(1)$$

$$= \kappa(1)\phi_3(1) \quad . \tag{9.15}$$

Similarly,

$$K_2(1)\phi_2(1) = \kappa(1)\phi_4(1) \tag{9.16}$$

$$K_3(1)\phi_3(1) = \kappa(1)\phi_1(1) \tag{9.17}$$

$$K_4(1)\phi_4(1) = \kappa(1)\phi_2(1) \quad . \tag{9.18}$$

Expressions (9.7), (9.8), (9.15)-(9.18) are used in (8.44); terms arising from the expansion of

$$\sum_{i=1}^{4} (J_i - K_i) \left(\phi_1(1)\phi_2(2)\phi_3(3)\phi_4(4) - \phi_1(1)\phi_4(2)\phi_3(3)\phi_2(4) \right.$$

$$\left. - \phi_3(1)\phi_2(2)\phi_1(3)\phi_4(4) + \phi_3(1)\phi_4(2)\phi_1(3)\phi_2(4) \right) \tag{9.19}$$

are shown in Appendix J.

At large R, where there is effectively no interaction, the molecular orbitals are simply the atomic orbitals, so that

$$\phi_1 = \chi_A \alpha \qquad (9.20a)$$

$$\phi_2 = \chi_A \beta \qquad (9.20b)$$

$$\phi_3 = \chi_B \alpha \qquad (9.20c)$$

$$\phi_4 = \chi_B \beta \qquad (9.20d)$$

then, using (9.20) in (8.8) and (8.9) to find J_i and $K_i\phi_i$, respectively, and integrating over spin and collecting terms:

$$J_1(1) = J_2(1) = \int \chi_A^2(r_x) r_{1x}^{-1} dr_x + 2 \int \chi_B^2(r_x) r_{1x}^{-1} dr_x \qquad (9.21)$$

$$J_3(1) = J_4(1) = 2 \int \chi_A^2(r_x) r_{1x}^{-1} dr_x + \int \chi_B^2(r_x) r_{1x}^{-1} dr_x \qquad (9.22)$$

$$K_1(1)\phi_1(1) = \left(\int \chi_A(x)\chi_B(x) r_{1x}^{-1} dr_x \right)\phi_3(1) = \kappa(1)\phi_3(1) \qquad (9.23)$$

$$K_2(1)\phi_2(1) = \kappa(1)\phi_4(1) \qquad (9.24)$$

$$K_3(1)\phi_3(1) = \kappa(1)\phi_1(1) \qquad (9.25)$$

$$K_4(1)\phi_4(1) = \kappa(1)\phi_2(1) \qquad . \qquad (9.26)$$

The terms arising from the expansion of (9.19) but where (9.20)-(9.26) are used for large R, are found in Appendix J.

At a late stage of this work, it was noted that a more general treatment is possible for all values of R by defining the orbitals:

$$\phi_1 = (c_{11}\chi_A + c_{12}\chi_B)\alpha \qquad\qquad (9.27a)$$

$$\phi_2 = (c_{11}\chi_A + c_{12}\chi_B)\beta \qquad\qquad (9.27b)$$

$$\phi_3 = (c_{21}\chi_A + c_{22}\chi_B)\alpha \qquad\qquad (9.27c)$$

$$\phi_4 = (c_{21}\chi_A + c_{22}\chi_B)\beta \qquad\qquad (9.27d)$$

and solving for J_i and $K_i\phi_i$ expressions. At small R, $c_{11} = c_{12}$ and $c_{21} = -c_{22}$. At large R, $c_{11} = c_{22} = 1$ and $c_{12} = c_{21} = 0$. At intermediate R, $c_{11} \neq c_{12} \neq c_{21} \neq c_{22} \neq 0$. Time did not permit exploring this approach.

(c)   Monte Carlo Algorithm

The variational energy as given by

$$E_{var} = \sum_i \varepsilon_i + \langle H_1 \rangle \Big/ \langle \psi^* \psi \rangle + \sum_i \langle \vec{\nabla}_i u \cdot \vec{\nabla}_i u \rangle \Big/ \langle \psi^* \psi \rangle \qquad (8.34)$$

is to be evaluated using a VMC calculation. This involves calculating two ratios of expected values, both with the same denominator $\langle \psi^* \psi \rangle$. From Part I of this work, the Monte Carlo estimate of E_{var} is

$$\hat{E}_{MC} = \sum_{i=1}^{N} \varepsilon_i + n^{-1} \sum_{\ell=1}^{n} \tilde{H}_1 \left[\underset{\sim}{r}^{(\ell)} \right] \Big/ n^{-1} \sum_{\ell=1}^{n} \tilde{S} \left[\underset{\sim}{r}^{(\ell)} \right]$$

$$+ \sum_{i=1}^{N} \left[n^{-1} \sum_{\ell=1}^{n} \tilde{G}_i \left[\underset{\sim}{r}^{(\ell)} \right] \Big/ n^{-1} \sum_{\ell=1}^{n} \tilde{S} \left[\underset{\sim}{r}^{(\ell)} \right] \right]$$

$$= \sum_{i=1}^{N} \varepsilon_i + \tilde{\tilde{H}}_1 / \tilde{\tilde{S}} + \sum_{i=1}^{N} \tilde{\tilde{G}}_i / \tilde{\tilde{S}}$$

$$= \sum_{i=1}^{N} \varepsilon_i + \hat{H}_1 + \sum_{i=1}^{N} \hat{G}_i \qquad (9.28)$$

where $\underset{\sim}{r}^{(\ell)}$ is the ℓ-th member of a set of n points in the configuration space of N electrons. The $\tilde{H}_1$, $\tilde{S}$, $\tilde{G}_i$ in (9.28) for the He_2 system under consideration where N = 4, are given by

$$\tilde{H}_1 = g\phi_{HP}U^2\left[\sum_{j>i}\sum_{i=1}^{4} r_{ij}^{-1} - \sum_{i=1}^{4}\left(J_i - K_i\right)\right]\det(\phi_1\phi_3)\det(\phi_2\phi_4)/\rho$$

(9.29)

$$\tilde{S} = g\phi_{HP}U^2\det(\phi_1\phi_3)\det(\phi_2\phi_4)/\rho$$

(9.30)

$$\tilde{G}_i = g\phi_{HP}U^2\left[\vec{\nabla}_i u\cdot\vec{\nabla}_i u\right]\det(\phi_1\phi_3)\det(\phi_2\phi_4)/\rho$$

(9.31)

where g is the Jacobian of transformation from Cartesian coordinates;

ρ is an arbitrary, normalized, everywhere positive probability distribution;

other symbols have been defined as ϕ_{HP} (8.41), $U = U\left(r_{ij}\right)$ (8.29), and

$$u = \sum_{j>i}\sum_{i} u\left(r_{ij}\right) \text{ in (8.29).}$$

As before, the estimates of the variances of the ratio estimators
$\hat{H}_1$ and $\hat{G}_i$ are

$$\hat{Var}(\hat{H}_1) = \left(n\bar{\tilde{S}}^2\right)^{-1}\left[\hat{Var}(\tilde{H}_1) - 2\hat{H}_1\hat{Cov}(\tilde{H}_1,\tilde{S}) + \hat{H}_1^2\hat{Var}(\tilde{S})\right]$$

(9.32)

and

$$\hat{Var}(\hat{G}_i) = \left(n\bar{\tilde{S}}^2\right)^{-1}\left[\hat{Var}(\tilde{G}_i) - 2\hat{G}_i\hat{Cov}(\tilde{G}_i,\tilde{S}) + \hat{G}_i^2\hat{Var}(\tilde{S})\right] .$$

(9.33)

The $\hat{E}_{MC}$ (9.28) involves the sum of ratio estimators and the

constant

$$\sum_{i=1}^{4}\varepsilon_i .$$

Now given a function $u(r_{ij})$ which is symmetrical in the coordinates, then let

$$\langle \tilde{G} \rangle \equiv \langle \tilde{G}_1 \rangle = \ldots = \langle \tilde{G}_4 \rangle \qquad (9.34)$$

and so

$$\sum_{i=1}^{4} \hat{G}_i \cong 4\hat{G} \cong 4\langle \tilde{G} \rangle / \langle \tilde{S} \rangle \qquad . \qquad (9.35)$$

Then substitution of (9.35) into (9.28) with N = 4 gives

$$\hat{E}_{MC} = \sum_{i=1}^{4} \varepsilon_i + \hat{H}_1 + 4\hat{G} \qquad . \qquad (9.36)$$

The variance of $\hat{E}_{MC}$ is then

$$\mathrm{Var}\left(\hat{E}_{MC}\right) = \mathrm{Var}(\hat{H}_1 + 4\hat{G})$$

$$= \mathrm{Var}\left[(\bar{\tilde{H}}_1 + 4\bar{\tilde{G}})/\bar{\tilde{S}}\right] \qquad . \qquad (9.37)$$

Appendix K gives a formula for the estimated variance of a sum of two ratio estimates having the same denominator. Using (K10) with $\hat{R} = \hat{H}_1$ and $\hat{R}' = 4\hat{G}$ so that $y = \tilde{H}$, $y' = 4\tilde{G}$ and $x = \tilde{S}$, then

$$\hat{\text{Var}}\left(\hat{E}_{MC}\right) \cong \left(n\bar{\bar{S}}^2\right)^{-1}\left\{\hat{\text{Var}}(\tilde{H}_1) + 2\hat{\text{Cov}}(\tilde{H}_1, 4\tilde{G}) + \hat{\text{Var}}(4\tilde{G})\right.$$

$$\left. - 2(\hat{H}_1 + 4\hat{G})\left[\hat{\text{Cov}}(\tilde{H}_1, \tilde{S}) + \hat{\text{Cov}}(4\tilde{G}, \tilde{S})\right] + (\tilde{H}_1 + 4\hat{G})^2\hat{\text{Var}}(\tilde{S})\right\} .$$

$$(9.38)$$

X. Monte Carlo estimate of ΔE_{INT}--Correlated Sampling

Consider the VMC calculation of the interaction energy, ΔE_{INT}; the Monte Carlo estimate of ΔE_{INT} (7.1), $\Delta \hat{E}_{MC}$ is given by

$$\Delta \hat{E}_{MC}(R) = \hat{E}_T(R) - \hat{E}_T\left(R_\infty\right) \quad . \tag{10.1}$$

where R_∞ is sufficiently large internuclear separation so that the two subsystems are not interacting; E_T is the total energy

$$E_T = E_{NN} + E_{HF} \tag{10.2}$$

where E_{NN} is the nuclear-nuclear repulsion term equivalent to

$$\sum_\alpha Z_\alpha^2 R^{-1}$$

(in this case $E_{NN} = 4R^{-1}$). The $\hat{E}_T$ is the Monte Carlo estimate of E_T,

$$\hat{E}_T = E_{NN} + \hat{E}_{MC} \tag{10.3}$$

where $\hat{E}_{MC}$ is given by (9.36). Because E_{NN} is a constant,

$$Var\left(\hat{E}_T\right) = Var\left(\hat{E}_{MC}\right) \quad . \tag{10.4}$$

If two independent runs are made to find $\hat{E}_T(R)$ and $\hat{E}_T(R_\infty)$, then the variance of $\Delta\hat{E}_{MC}(R)$ is the sum of the variances of the two estimates:

$$\text{Var}\left[\Delta\hat{E}_{MC}(R)\right] = \text{Var}\left[\hat{E}_{MC}(R)\right] + \text{Var}\left[\hat{E}_{MC}(R_\infty)\right] \quad . \tag{10.5}$$

This is likely to be significant, especially since $\Delta\hat{E}_{MC}(R)$ is much less than either $\hat{E}_T(R)$ or $\hat{E}_T(R_\infty)$. A powerful method of variance reduction is Correlated sampling [62] in which the same set of random numbers are used to calculate both $\hat{E}_{MC}(R)$ and $\hat{E}_{MC}(R_\infty)$. The objective is to create a high positive correlation between the samples which causes

$$\text{Var}\left[\Delta\hat{E}_{MC}(R)\right] = \text{Var}\left[\hat{E}_{MC}(R)\right] + \text{Var}\left[\hat{E}_{MC}(R_\infty)\right] - 2\text{Cov}\left[\hat{E}_{MC}(R),\hat{E}_{MC}(R_\infty)\right]$$

$$\tag{10.6}$$

to be much less than $\text{Var}\left(\Delta\hat{E}_{MC}(R)\right)$ given by (10.5), where independent samples cause the covariance term in (10.6) to be zero.

In this situation, the easiest method of creating positive correlation is to use the same set of random numbers $\underline{p}^{(i)}$ to generate configuration points $\underline{r}(R)^{(i)}$ and $\underline{r}\left(R_\infty\right)^{(i)}$ used to calculate $\hat{E}_{MC}(R)$ and $\hat{E}_{MC}\left(R_\infty\right)$, respectively. If the Importance sampling functions used to generate $\underline{r}(R)^{(i)}$ and $\underline{r}\left(R_\infty\right)^{(i)}$ are similar, then the two samples of configuration points will be correlated and so will $\hat{E}_{MC}(R)$ and $\hat{E}_{MC}\left(R_\infty\right)$.

The one complication added by the use of Correlated sampling is in calculating the $\text{Cov}\left(\hat{E}_{MC}(R),\hat{E}_{MC}\left(R_\infty\right)\right)$ term in (10.6). While in principle

this is easily found, in practice it entails that either (i) simultaneous simulations be run so that a cumulative calculation of the covariance can be made, or (ii) separate simulations be run but storage of $E_{MC}^{(i)}(R)$ and $E_{MC}^{(i)}(R_\infty)$, i = 1, ..., n, allows a later computation of the covariance. In the former case, a restriction on computer time may be a problem and in the latter, large amounts of storage space are necessary.

From (9.36) and (10.1),

$$\Delta \hat{E}_{MC} = \sum_{i=1}^{4} \left(\varepsilon_i(R) - \varepsilon_i(R_\infty) \right) + \left(\hat{H}_1(R) + 4\hat{G}(R) \right) - \left(\hat{H}_1(R_\infty) + 4\hat{G}(R_\infty) \right)$$

(10.7)

which involves differences between ratio estimates with different denominators. Appendix K provides a formula (K12) which can be used to estimate $\text{Var}(\hat{R} - \hat{R}')$ where

$$\hat{R} = \left(\bar{\bar{H}}_1(R) + 4\bar{\bar{G}}(R) \right) / \bar{\bar{S}}(R)$$

(10.8)

and

$$\hat{R}' = \left(\bar{\bar{H}}_1(R_\infty) + 4\bar{\bar{G}}(R_\infty) \right) / \bar{\bar{S}}(R_\infty)$$

(10.9)

and (K10) is used to find both $\hat{\text{Var}}(\hat{R})$ and $\hat{\text{Var}}(\hat{R}')$. Equation (K13) is used to find $\hat{\text{Cov}}(\hat{R}, \hat{R}')$ where $y = \tilde{H}_1(R) + 4\tilde{G}(R)$, $x = \tilde{S}(R)$, $y' = \tilde{H}_1(R_\infty) + 4\tilde{G}(R_\infty)$ and $x' = \tilde{S}(R_\infty)$.

XI. Choice of Correlated Function--Simulations of SCF Results

--Optimization of $\psi(\underline{x})$-VMC Results

(a) Choice of $U(\underline{r}_{ij})$

The wavefunction chosen for the VMC calculations is described by

(8.43) where the spin orbitals ϕ_i for small and large internuclear

separations are given by (9.5) and (9.20), respectively; atomic orbitals

χ_A and χ_B are given by (9.4) with $\zeta = 1.69$, the optimized orbital

exponent for an STO describing He.

For these initial studies, two simple correlation functions,

$U(\underline{r}_{ij})$ (8.29) were chosen on the basis of other calculations found in the

literature. These are

$$\exp\left[-\sum_{j>i}\sum_i u\left(r_{ij}\right)\right] = \exp\left[-\sum_{j>i}\sum_i ar_{ij}\Big/\left(1 + br_{ij}\right)\right] \qquad (11.1)$$

and

$$\exp\left[-\sum_{j>i}\sum_i u(r_{ij})\right] = \exp\left[-\sum_{j>i}\sum_i\left[1 - \exp(-cr_{ij})\right]\Big/c^2 r_{ij}\right] \qquad (11.2)$$

where a, b, c are variational parameters. Equation (11.1) is an example

of a Jastrow function [63]; Moskowitz and Kalos [26], Handy [64],

Reynolds et al. [25] have employed functions of this type in atomic

and molecular (chemically bound systems) calculations. This Padé form

obeys the electron-electron cusp condition, which requires $u(r_{ij})$ to be

linear in r_{ij} at small r_{ij}, and the requirement that $u(r_{ij})$ asymptotically

approach a constant and be of the order $O(r_{ij}^{-1})$ so that the wavefunction

factors at large r_{ij} [25, 65].

The second $U(\underset{\sim}{r}_{ij})$ (11.2) has been used by Gibbs and Dunn [6] in a

calculation of C^{2+}. It also has the aforementioned properties as (11.1)

at small and large r_{ij}.* Other functional forms for the correlation

function [66] have not been examined by this work.

Appendix L contains expressions for $\vec{\nabla}_i u \cdot \vec{\nabla}_i u$ for both (11.1) and

(11.2).

* After calculations using (11.2) had been made, it was noted that

(11.2) is not unitless. Further investigation into references quoted by

Gibbs and Dunn [67, 68] lead to speculation that a misprint was made in

[6] and that the functional form of (11.2) would more properly involve

something such as

$$\sum_{j>i} \sum_i a \left[1 - \exp(-cr_{ij}) \right] r_{ij}^{-1} \quad .$$

The use of an incorrect correlation function would, of course, invalidate

the results found in (XIc) and in Tables III and IV.

(b) Computer Program--Simulation of SCF Results for He_2

The HF-SCF program HONDO [59] was run on the Burroughs 6700 in order to obtain orbital energies, ε_i, molecular coefficients c_1 and c_2 (9.5) and the SCF results for He_2 at various internuclear separations between 1 au and 14 au. The calculations employ an STO-6G minimal basis set (a computational check was performed to ensure that the particular combination of Gaussians used by HONDO mimic an STO with $\zeta = 1.69$).

The range over which χ_A and χ_B contribute equally to ϕ_i (9.5) lies below about 9 au. The value of R_∞ was taken at 14 au, beyond which the atoms were effectively non-interacting ((9.20) could be employed).

Points in configuration space used to calculate $\hat{H}_1$ (9.28) and $\hat{G}$ (9.36) are chosen using Importance sampling in prolate spheroidal coordinates. The probability inversion technique is employed; the method is described in Appendix M.

Computer programs were written in FORTRAN IV to run on the VAX-11/780 in double precision which calculate $\hat{E}_{MC}$ (9.36), $\hat{Var}(\hat{E}_{MC})$ (9.38) and other quantities of interest at small and large R. The $\hat{G}$ in (9.36) is computed as $\hat{G}_1$; no attempt was made to compare the $\hat{G}_i$, i = 1, ..., 4, values. Intermediate ranges (9 au < R < 14 au) where (9.20) hold were not considered. The sample means, variances and covariances are found cumulatively using an efficients moments routine [69]. Standard deviations are taken as the square roots of the estimated variances.

The programs were first used to find the Monte Carlo estimate of E_{HF}, $\hat{E}_{HF}$, that is, where $U(\underset{\sim}{r}_{ij})$ is set to unity so that $\psi(\underline{x}) = \Phi(\underline{x})$ and

$$\hat{E}_{MC}\left[U(\underset{\sim}{r}_{ij}) = 1\right] = \hat{E}_{HF} = \sum_{i=1}^{4} \varepsilon_i + \hat{H}_1 \qquad (11.3)$$

Table II contains values of ε_i, c_1 and c_2, E_{HF}, $-\frac{1}{2}\langle V_{HF}\rangle$,

$$\left\langle \sum_{j>i} \sum_{i} r_{ij}^{-1}\right\rangle$$

from HONDO and the Monte Carlo ratio estimates, $\hat{E}_{HF}$, $\hat{H}_1$, $\hat{V}_{MC}$ and $\Sigma\hat{r}_{ij}^{-1}$ of E_{HF},

$$\left\langle \sum_{j>i} \sum_{i} r_{ij}^{-1} - V_{HF}\right\rangle, \langle V_{HF}\rangle \text{ and } \left\langle \sum_{j>i} \sum_{i} r_{ij}^{-1}\right\rangle$$

respectively, given an independent sample of 10^4 points. Also noted are standard deviations of the ratio estimates (analogous equations to (9.32) and (9.33) are used to estimate $\widehat{Var}^{\frac{1}{2}}(\hat{V}_{MC})$ and $\widehat{Var}^{\frac{1}{2}}\left(\Sigma\hat{r}_{ij}^{-1}\right)$).

The advantage of using the combined estimate $\hat{H}_1$ over a separate calculation of $\hat{V}_{MC}$ and $\Sigma\hat{r}_{ij}^{-1}$ is that it is immediately apparent that positive correlation between elements used to calculate $\hat{V}_{MC}$ and $\Sigma\hat{r}_{ij}^{-1}$ has reduced the variance of $\hat{H}_1$. That is, the variance of $\hat{H}_1$ is not simply the sum of the variances of $\hat{V}_{MC}$ and $\Sigma\hat{r}_{ij}^{-1}$. Although $\hat{H}_1$ was treated as one ratio estimator, appropriate use of (K10) could have been used to find $\widehat{Var}(\hat{H}_1)$ if the difference of the two ratio estimates $\hat{V}_{MC}$

TABLE II   HONDO and MONTE CARLO Results $(n=10^4)$ for SCF Calculation of ${}^1\Sigma_g^+He_2$ [a].

R [b]	c_1 [c]	c_2 [c]	$\varepsilon_1 = \varepsilon_2$ [d]	$\varepsilon_3 = \varepsilon_4$ [d]	E_{HF} [e]	$\widetilde{E}_{HF}$ [f]
1.0	0.5468442	1.234709	-1.48786373	-0.12353766	-8.253	-8.17 ± .03
3.0	0.6766807	0.7420445	-0.94862522	-0.83608068	-7.013	-7.03 ± .02
5.6	0.7060456	0.7081728	-0.89676836	-0.89326402	-6.407	-6.44 ± .02
6.0	0.7065173	0.7076977	-0.89601513	-0.89402505	-6.359	-6.39 ± .02
7.0	0.7069819	0.7072317	-0.89525308	-0.89479082	-6.264	-6.29 ± .02
9.0	0.7071027	0.7071108	-0.89503196	-0.89501215	-6.137	-6.14 ± .02
14.0	non-interacting AO's		-0.89502206	-0.89502206	-5.978	-6.06 ± .03

R [b]	$-\frac{1}{2}\langle V_{HF}\rangle$ [g]	$\widehat{H}_1$ [h]	$\widehat{V}_{MC}$ [i]	$\overline{\Sigma r_{ij}^{-1}}$ [j]
1.0	-5.030	-4.95 ± .03	9.97 ± .04	5.02 ± .03
3.0	-3.444	-3.46 ± .02	6.89 ± .01	3.43 ± .02
5.6	-2.827	-2.86 ± .02	5.66 ± .01	2.80 ± .02
6.0	-2.779	-2.81 ± .02	5.56 ± .01	2.75 ± .02
7.0	-2.684	-2.71 ± .02	5.37 ± .01	2.66 ± .02
9.0	-2.557	-2.59 ± .02	5.15 ± .01	2.56 ± .02
14.0	-2.398	-2.48 ± .03	4.74 ± .03	2.27 ± .04

[a] $\psi(x)$ is the SCF wavefunction for He_2; HONDO uses STO-6G to mimic STO with $\zeta=1.69$; the Monte Carlo program uses $\psi(x) = \det(\phi_1\phi_3)\det(\phi_2\phi_4)$ with ϕ_i defined by (9.5) for $R\leq 9au$ and by (9.20) for $R=14.0au$; $\zeta=1.69$ in (9.4). All units in Table II are in au.

[b] Internuclear separation.

[c] Coefficients of molecular orbitals as in (9.5).

[d] Orbital energies of (9.5).

[e] E_{HF} from HONDO.

[f] $\widetilde{E}_{MC}$ as in (9.36) with $\widehat{G} = 0$. Hence $\widehat{E}_{MC} = \widetilde{E}_{HF}$ and $\mathrm{Var}^{\frac{1}{2}}(\widehat{E}_{HF}) = \mathrm{Var}^{\frac{1}{2}}(\widehat{H}_1)$.

[h] $\widehat{H}_1$ as in (9.28); $\widehat{H}_1$ estimates $-\frac{1}{2}\langle V_{HF}\rangle$.

[g] $-\frac{1}{2}\langle V_{HF}\rangle = \langle\sum_{j>i}\sum_i r_{ij}^{-1}\rangle$ (8.27).

[i] $\widehat{V}_{MC}$ estimates $\langle V_{HF}\rangle$ (8.22).

[j] $\overline{\Sigma r_{ij}^{-1}}$ estimates $\langle\sum_{j>i}\sum_i r_{ij}^{-1}\rangle$ (8.27).

and $\Sigma \hat{r}_{ij}^{-1}$ had been employed.

All the SCF results lie within three standard deviations of the Monte Carlo estimates; the three standard deviation confidence intervals of $\bar{H}_1$ and $-\Sigma \hat{r}_{ij}^{-1}$ overlap. Figure II is a plot of the Hartree-Fock energy, E_{HF} and the Monte Carlo estimate $\hat{E}_{HF}$ versus R, along with $3\widehat{Var}^{\frac{1}{2}}(\hat{E}_{HF})$ confidence intervals.

Figure II. Monte Carlo Estimate $\hat{E}_{HF}$ of the

Hartree-Fock Energy <u>versus</u> Internuclear

Separation R for He_2
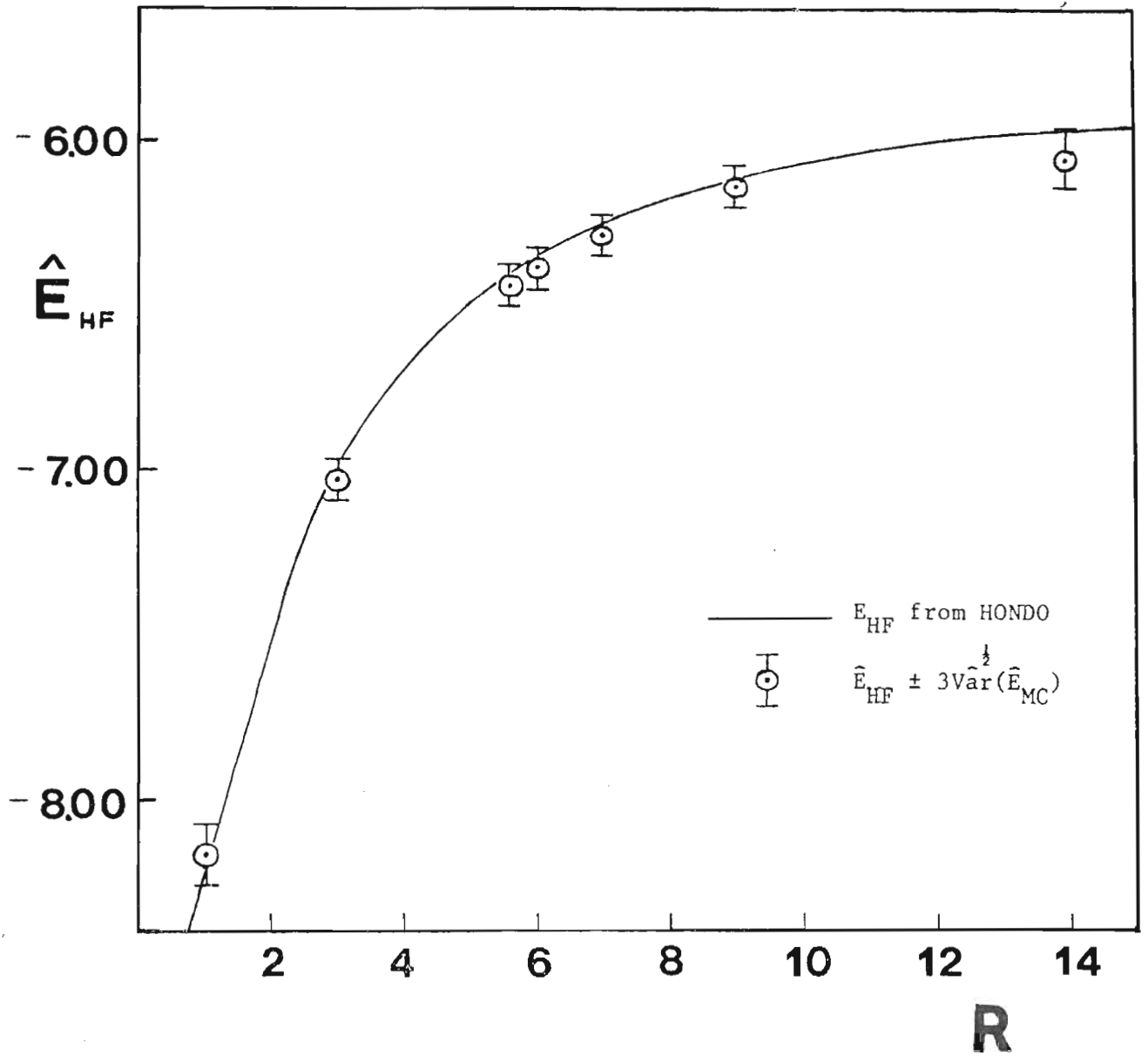
(c) Optimization of $\psi(\underline{x})$--Calculation of $\Delta\hat{\bar{E}}_{MC}$

The accepted minimum in the energy versus internuclear separation curve falls at about 5.6 au [50-55] which is where Coldwell and Lowther [7] decided to optimize their wavefunction. Using the results from Part I of this work, the variational parameters in $U(\underline{r}_{ij})$ of the explicitly correlated wavefunction were to be optimized by finding the minimum $\hat{\bar{E}}_{MC}$.

Parameter a in the first correlation function was set to 0.5* ($\psi(\underline{x})$ then rigorously obeys the cusp condition) and optimization with respect to b was attempted using a sample of 10^3 configuration points and subroutine VA04A [39]. This was not successful as no minimum was found: b was sent to increasingly large values. This raised the question as to the suitability of (11.1) for this calculation, which is addressed in the discussion section, Chapter XII.

It was possible to optimize the wavefunction with respect to parameter c of (11.2). Table III contains values of $\hat{c}^*$ (estimate of optimal value of c) and corresponding values of $\hat{\bar{E}}_{MC}$ (with one standard deviation) and $\hat{E}_T$ for 10 different samples of 10^3. The average $\bar{\hat{c}}^*$ and $\bar{\hat{E}}_{MC}$ and $\bar{\hat{E}}_T$ with standard deviations are also shown.

It should be noted that $\hat{E}_{MC}$ is not very sensitive to the value of c; a change in c in the vicinity of $\hat{c}^*$ by one unit in either direction only affects $\hat{E}_{MC}$ in the third decimal place, hence the 20% variation in

*Parameter a should have in fact been set to -0.5 in order to obey the cusp condition.

TABLE III Optimization of a One-Parameter Correlated Wavefunction for $^1\Sigma_g^+$ He$_2$.[a]

Sample[b]	$\hat{c}^*$[c]	$\hat{E}_T$[d]	$\hat{E}_{MC}$[e]
1	4.83	$-5.77 \pm .05$	$-6.48 \pm .05$
2	4.35	$-5.71 \pm .06$	$-6.42 \pm .06$
3	4.83	$-5.76 \pm .05$	$-6.47 \pm .05$
4	5.24	$-5.71 \pm .06$	$-6.43 \pm .06$
5	4.18	$-5.60 \pm .09$	$-6.32 \pm .09$
6	6.50	$-5.69 \pm .07$	$-6.40 \pm .07$
7	4.06	$-5.73 \pm .05$	$-6.44 \pm .05$
8	4.16	$-5.65 \pm .08$	$-6.36 \pm .08$
9	4.76	$-5.74 \pm .05$	$-6.46 \pm .05$
10	6.57	$-5.69 \pm .07$	$-6.41 \pm .07$

Averages Computed from the Above Samples:

$\overline{\hat{c}}^* = 5.0 \pm .9 \qquad \overline{\hat{E}}_T = -5.71 \pm .05 \qquad \overline{\hat{E}}_{MC} = -6.42 \pm .05$

[a] Trial wavefunction $\psi(\underline{x})$ is given by (8.43); $U(\underline{r}_{ij})$ from (11.2); ϕ_i from (9.5) with $\zeta = 1.69$; c_1, c_2, ε_i for R = 5.6 au in Table II.

[b] Separate samples of 10^3 Monte Carlo points.

[c] $\hat{c}^*$ estimates optimum c in $U(\underline{r}_{ij})$ (11.2) which gives lowest E_{var} (8.34).

[d] $\hat{E}_T$ given by (10.3).

[e] $\hat{E}_{MC}$ given by (9.36).

$\hat{c}^*$ values as compared to the 1% variation in $\bar{\hat{E}}_{MC}$. This means that it is difficult to "miss" the optimal value of c, which is important since only a relatively crude estimate of c^* is being made using small samples. Larger simulations were not used, as to optimize one sample takes about $2\frac{1}{2}$ hours of processor time.

Two larger simulations with $n = 5 \times 10^4$ were run separately but using the same set of random numbers setting $c = \bar{\hat{c}}^* = 5.0$ at $R = R_e = 5.6$ au and at $R = R_\infty = 14.0$ au. Table IV contains the results, including $\hat{E}_{MC}$, $\hat{E}_T$, $\hat{V}_{MC}$, $\Sigma\hat{r}_{ij}^{-1}$, $\bar{H}_1$, estimates $4\hat{T}_2$ and $4\hat{T}_3$ of the two- and three-body contributions to $4\hat{G}$, respectively, $4\hat{G}$ and the processor times for the runs.

Note that two-body contributions to $4\hat{G}$ predominate, as found by Gibbs and Dunn [6]. Consider that the vdW complex at R_e is still very much like two separate atoms, hence interactions between the two electrons centered on the same nucleus will be stronger than either two- or three-body interactions involving electrons on different atoms. As R increases, the magnitudes of $4\hat{T}_2$ and especially $4\hat{T}_3$ decrease, as expected: A and B interaction decreases with increasing separation.

Comparison of $\hat{E}_T$ with those values obtained by other workers [51, 54, 70] shows that the as yet crude estimates are within one standard deviation of the simple five-configuration MCSCF calculation. This is encouraging but should be used only as a guideline for the suitability of $\psi(\underline{x})$ for a qualitative calculation of ΔE_{INT}: recall that in such variational calculations that theoreticians typically assume that

TABLE IV $\hat{E}_{MC}$ for $^1\Sigma_g^+ He_2$ using an Explicitly Correlated Wavefunction. (n = 5 x 10^4). [a]

	R = 5.6	R = 14.0
$\hat{V}_{MC}$ [b]	5.629 ± .006	4.76 ± .01
$\hat{\Sigma r_{ij}^{-1}}$ [c]	2.751 ± .009	2.31 ± .02
$\hat{H}_1$ [d]	-2.878 ± .008	-2.45 ± .01
$4\hat{T}_2$ [e]	0.021 ± .003	0.012 ± .001
$4\hat{T}_3$ [f]	4 ± 2 (-7)	3 ± 1 (-7)
$4\hat{G}$ [g]	0.021 ± .003	0.012 ± .001
$\hat{E}_{MC}$ [h]	-6.44 ± .01	-6.02 ± .02
$\hat{E}_T$ [i]	-5.72 ± .01	-5.74 ± .02
CPU-time [j]	8:03	8:23
E_T	-5.72339231 [k]	R_∞ -5.72335924 [k]
	-5.79530789 [l]	R_∞ -5.79534108 [l]
	-5.80416168 [m]	R_{20} -5.80412829 [m]

[a] $\psi(\underset{\sim}{x})$ is (8.43); $U(\underset{\sim}{r}_{ij})$ is (11.2) with c = 5.0; ϕ_i from (9.5) with ζ = 1.69 c_1, c_2, ε_i for R = 5.6 au in Table II; ϕ_i from (9.20) with ζ = 1.69 ϕ_i for R = 14.0 in Table II.

[b] $\tilde{V}_{MC}$ estimates $\langle V_{HF} \rangle$ (8.22).

[c] $\hat{\Sigma r_{ij}^{-1}}$ estimates $\langle \underset{j>i}{\sum}\underset{i}{\sum} r_{ij}^{-1} \rangle$ (8.27).

[d] $\hat{H}_1$ as in (9.28).

TABLE IV continued...

^{e} $\hat{T}_2$ estimates two-body contributions to $\langle \vec{\nabla}u \cdot \vec{\nabla}u \rangle / \langle \psi * \psi \rangle$; see Appendix L, (L14).

^{f} $\hat{T}_3$ estimates three-body contributions to $\langle \vec{\nabla}u \cdot \vec{\nabla}u \rangle / \langle \psi * \psi \rangle$; see Appendix L, (L15).

^{g} $\hat{G}$ estimates $\langle \vec{\nabla}u \cdot \vec{\nabla}u \rangle / \langle \psi * \psi \rangle$; see (9.31), (9.34), (9.35), (L16).

^{h} $\hat{E}_{MC}$ from (9.36).

^{i} $\hat{E}_T$ from (10.3).

^{j} Computer processor time -- hours:minutes.

^{k} 5 - configuration MCSCF calculation [70].

^{l} 20 - configuration MCSCF calculation [51].

^{m} CEPA - PNO calculation [54]; R_{20} is R = 20 au.

the absolute error in E_{var} remains constant with R. The error in E_{var} is usually greater than the well-depth itself (Ch. VII).

Note that $\hat{E}_T$ values are not converging towards the more accurate twenty-configuration MCSCF or CEPA-PNO result. It is apparent that this simple correlation function cannot recover all of the correlation energy.

The errors in the estimates $\hat{E}_T$ are too large to calculate an interaction energy. Even if correlated sampling had been used, it is doubtful whether the reduction in variance would increase the accuracy in the difference beyond the second decimal place. In order to obtain a sample estimate of $\hat{E}_T$ accurate to the fifth decimal, it is estimated that a sample on the order of 10^{10} would be needed. This would be an extremely lengthy calculation given the processor time for a 5×10^4 calculation. It is felt that at this point, time would be more wisely invested in making modifications to the existing programs in order to make them more efficient. Such modifications are outlined below.

(i) The major time consuming element in the calculation involves the probability inversion which finds ξ_i and η_i for the configuration points $(\underline{r}^{(i)})$, $i = 1, \ldots, n$. Other methods of generating random variables [29] should be investigated on the basis of efficiency.

(ii) Simultaneous calculations at $R = R_e$ and $R = R_\infty$ should be run, incorporating Correlated sampling (Ch. X), so that advantage is taken of the variance reduction offered. In addition, other methods of variance reduction, such as the use of control variates, should be exploited. A reduction of variance of close to 50% has been obtained in Monte Carlo simulations using this technique [38, 71].

(iii) An alternate method of Importance sampling for short and long R may be more appropriate. It may be too much to expect one technique to encompass the whole range of R (Appendix M: Note the different transformations necessary in order to find η at short and long R.)

(iv) Incorporate Equations (9.27) for ϕ_i.

(v) General streamlining of FORTRAN code is desirable.

(vi) In the light of the footnote in (XIa), the corrected form of (11.2) should be used in the simulation.

(vii) In addition, the parameters were only optimized at R_e = 5.6 au. In practice they should be optimized at both internuclear separations.

XII. Discussion and Summary

The results permit only a short discussion on the suitability of correlation functions (11.1) and (11.2). It was found that $\psi(\underline{x})$ could not be optimized with respect to one variational parameter, b, in (11.1), while optimization with respect to c in (11.2) was possible. Examination of the two functions shows that they behave quite differently as functions of r_{ij}. Since b in (11.1) is sent to very large values, the correlation function tends towards a constant. This may mean that large r_{ij} values become too important in the optimization of this function. The relaxation of the condition, a = 0.5,* and optimizing with respect to a and b may allow $\psi(\underline{x})$ using (11.1) to be optimized.

In vdW molecules most of the correlation energy arises from interactions between electrons on different atoms and the intra-atomic correlation energy remains essentially constant. Both (11.1) and (11.2) are expected to deal with both inter- and intra-atomic correlation energy. A $U(\underline{r}_{ij})$ which deals with the two effects separately would be more versatile.

The Basis Set Extension Effect was also not accounted for in these calculations. Some technique, such as the use of ghost orbitals, should be incorporated in order to avoid a fortuitous minimum.

In summary, in Part I of this work, theoretical derivations show that in VMC calculations which employ the ratio estimate of E_{var} (4.8)

---

* See footnote on p. 75.

have a bias of the estimate which is negligible for large samples (4.9).
Aside from the bias terms, the expected value of the estimate is the
variational energy (the exact energy plus terms second order in smallness
parameters (4.7)). The variance of the energy estimate is also second
order (4.13).

Given an infinite simple random sample, optimum variational
parameters obtained by minimizing the variance of the ratio estimate,
$\underset{\sim}{\beta}^{**}$ (4.15), are not identical to those which minimize the variational
energy, $\underset{\sim}{\beta}^{*}$ (4.14). This provides a theoretical explanation for the need
to rescale after estimates of $\underset{\sim}{\beta}^{**}$ are obtained.

For normalized wavefunctions, the VMC energy estimate (4.17) is
unbiased; it has E_{var} for its expected value (4.18). However, for large
samples, the biased ratio estimator is preferred because it has a smaller
variance due to the correlation of two random variables, H (4.3) and S
(4.4). The theoretical result given in (4.19) shows that the unbiased
estimate has a variance which is zeroth order in smallness parameters.

A numerical example using Importance sampling and a Hylleraas
variational wavefunction for $1\ ^{1}S$ He is given. The results indicate that
there is no advantage to using minimization of the variance as a
parameter optimization scheme.

The diffusion Monte Carlo energy estimate of Reynolds et al. [25]
has negligible bias for large t (5.14). (It is assumed that the nodes
of the exact wavefunction correspond exactly to those of the guiding
function.) The estimate yields the exact energy in the limit of

infinite t. However, the variance of the energy estimate is second order in smallness parameters, even at infinite t (5.15c). This provides a theoretical justification for the empirical observation that the precision of DMC calculations is strongly dependent on the choice of the guiding function.

In Part II, it is shown that a partitioned form of the Hamiltonian (8.24) can be used in conjunction with an explicitly correlated wavefunction (8.28), which allows E_{var} to be expressed by a combination of HF-SCF results and terms involving the gradient of the correlation function (8.34).

Equations are derived which allow the Hartree-Fock potential, V_{HF} (8.14), expressed in terms of coulomb and exchange operators, to be applied to an explicitly correlated wavefunction for the vdW molecule, He_2. A method of Importance sampling in prolate spheroidal coordinates is derived. A VMC calculation is performed for He_2, firstly for the HF-SCF wavefunction and secondly for an explicitly correlated wavefunction. The latter employs results of Part I to optimize $\psi(\underset{\sim}{x})$ for one form of $U(\underset{\sim}{r}_{ij})$ (11.2). A wavefunction using another correlation function (11.1) could not be optimized. Although the Monte Carlo estimates at short and long distances agree within statistical error with representative (qualitative) MCSCF results, it is not possible to estimate ΔE_{INT} because of insufficiently accurate values of $\hat{E}_{MC}$. Modifications which could allow the computer program to run more efficiently are discussed.

Appendix A. The Central Limit Theorem (CLT) of Probability

The Central Limit Theorem [34] states that, given n independent variables $\{x_i\}$ with means $\{m_i\}$ and standard deviations $\{\sigma_i\}$, the sum

$$x = \sum_{i=1}^{n} x_i \qquad \text{(A1)}$$

is asymptotically normally distributed

$$p(x) \sim N(m,\sigma) \qquad \text{(A2)}$$

with mean m and variance σ^2 where

$$m = \sum_{i=1}^{n} m_i \qquad \text{(A3)}$$

and

$$\sigma^2 = \sum_{i=1}^{n} \sigma_i^{\ 2} \qquad . \qquad \text{(A4)}$$

The distributions of $\{x_i\}$ do not have to be normal and are subject only to certain very general conditions. This theorem was first stated by LaPlace [72] and involves a fairly complicated proof.

An important extension is that if the $\{x_i\}$ are identically distributed so that

$$m_1 = m_2 = \ldots = m_n = m_c \tag{A5}$$

and

$$\sigma_1 = \sigma_2 = \ldots = \sigma_n = \sigma_c \quad , \tag{A6}$$

then the sample mean $\bar{x}$

$$\bar{x} = n^{-1} \sum_{i=1}^{n} x_i \tag{A7}$$

is also asymptotically normal with mean m_c and variance σ_c^2/n. For the CLT to hold in this case, it is sufficient to assume that the distribution of $\{x_i\}$ has a finite second moment, that is, σ_c^2 exists and is finite.

A random variable x' is said to be normally distributed with mean m and variance σ^2 if the distribution function of x' follows

$$p(x) = (2\pi\sigma^2)^{-\frac{1}{2}} \exp\left(-(x - m)^2/2\sigma^2\right) \quad . \tag{A8}$$

The probability, P, that x' differs from m in either direction by more than $\lambda\sigma$ is equal to the combined areas under the curve $p(x)$ beyond $x = m + \lambda\sigma$ and below $x = m - \lambda\sigma$:

$$P\left(|x' - m| > \lambda\sigma\right) = 2(2\pi\sigma^2)^{-\frac{1}{2}} \int_{\lambda\sigma}^{\infty} \exp\left(-(x - m)^2/2\sigma^2\right) dx \tag{A9}$$

where the symmetry of $p(x)$ about $x = m$ has been employed. Equation (A9) reduces to

$$P\left(\left|x' - m\right| > \lambda\sigma\right) = (2/\pi)^{\frac{1}{2}} \int_{\lambda}^{\infty} \exp(-x^2/2)\,dx \qquad . \tag{A10}$$

Thus it is stated in the text (3.12), that

$$P\left(\left|x' - m\right| < 3\sigma\right) = 1 - P\left(\left|x' - m\right| > 3\sigma\right) \cong 0.997 \tag{A11}$$

where (A10) is solved with $\lambda = 3$. Similarly,

$$P\left(\left|x' - m\right| < \sigma\right) \cong 0.683 \qquad . \tag{A12}$$

Appendix B. Derivations of (4.7) and (4.13)

Given a trial wavefunction ψ_t (4.1) and expressions for $\langle H \rangle$ (4.3) and $\langle S \rangle$ (4.4) as given in the main body of this work, then E_{var} (4.2) is

$$E_{var} = \left(E_o + \sum_{k=1}^{\infty} \lambda_k^2 E_k \right) \left(1 + \sum_{k=1}^{\infty} \lambda_k^2 \right)^{-1} \quad . \tag{B1}$$

Assuming that

$$1 > \sum_k \lambda_k^2 \quad ,$$

$\langle S \rangle^{-1}$ can be expressed as a binomial series so that

$$E_{var} = E_o + \sum_k \lambda_k^2 (E_k - E_o) + 0(\lambda^4) \quad . \tag{4.7}$$

In this and all formulae which follow, the summations run over excited states.

In order to express the variance of $\hat{E}_{MC}$ in terms of λ, it is necessary to obtain expressions for Var(H), Cov(H,S) and Var(S). From (4.1), (4.3) and

$$Var(H) \equiv \langle H^2 \rangle - \langle H \rangle^2 \quad , \tag{B2}$$

explicitly to third order,

$$\text{Var}(H) = E_o^2 \langle \phi_o^4 \rangle + 2\sum_k \lambda_k E_o (E_o + E_k) \langle \phi_o^3 \phi_k \rangle$$

$$+ \sum_{k,\ell} \lambda_k \lambda_\ell (E_o^2 + 4E_o E_k + E_k E_\ell) \langle \phi_o^2 \phi_k \phi_\ell \rangle$$

$$+ 2\sum_{k,\ell,m} \lambda_k \lambda_\ell \lambda_m E_k (E_o + E_\ell) \langle \phi_o \phi_k \phi_\ell \phi_m \rangle$$

$$- (E_o^2 + 2\sum_k \lambda_k^2 E_o E_k) + 0(\lambda^4) \quad . \tag{B3}$$

The last term in (B3) corresponds to $\langle H \rangle^2$.

Similarly,

$$\text{Var}(S) \equiv \langle S^2 \rangle - \langle S \rangle^2 \tag{B4}$$

$$= \langle \phi_o^4 \rangle + 4\sum_k \lambda_k \langle \phi_o^3 \phi_k \rangle + 6\sum_{k,\ell} \lambda_k \lambda_\ell \langle \phi_o^2 \phi_k \phi_\ell \rangle$$

$$+ 4\sum_{k,\ell,m} \lambda_k \lambda_\ell \lambda_m \langle \phi_o \phi_k \phi_\ell \phi_m \rangle - (1 + 2\sum_k \lambda_k^2) + 0(\lambda^4) \quad , \tag{B5}$$

and

$$\text{Cov}(H,S) \equiv \langle HS \rangle - \langle H \rangle \langle S \rangle \tag{B6}$$

$$= E_o \langle \phi_o^4 \rangle + \sum_k \lambda_k (3E_o + E_k) \langle \phi_o^3 \phi_k \rangle$$

$$+ 3\sum_{k,\ell} \lambda_k \lambda_\ell (E_k + E_o) \langle \phi_o^2 \phi_k \phi_\ell \rangle$$

$$+ \sum_{k,\ell,m} \lambda_k \lambda_\ell \lambda_m (E_o + 3E_k) \langle \phi_o \phi_k \phi_\ell \phi_m \rangle$$

$$- \left(E_o + \sum_k \lambda_k^2 (E_o + E_k) \right) + 0(\lambda^4) \qquad . \tag{B7}$$

The variance of the ratio estimate, $\text{Var}(\hat{E}_{MC})$, is given by,

$$\text{Var}(\hat{E}_{MC}) = (n\langle S \rangle^2)^{-1} \left(\text{Var}(H) - 2E_{var} \text{Cov}(H,S) + E_{var}^2 \text{Var}(S) \right) + 0(n^{-2}) \quad , \tag{4.10}$$

where $0(n^{-2})$ reflects the bias. After substituting (4.6), (4.7), (B3), (B5) and (B7) into (4.10), all zero and first-order terms in λ drop out, leaving

$$\text{Var}(\hat{E}_{MC}) = n^{-1} \left(1 + \sum_k \lambda_k^2 \right)^{-2} \left[\sum_{k,\ell} \lambda_k \lambda_\ell (E_k - E_o)(E_\ell - E_o) \langle \phi_o^2 \phi_k \phi_\ell \rangle \right.$$

$$- 2\sum_{k,\ell} \lambda_k^2 \lambda_\ell (E_k - E_o)(E_\ell - E_o) \langle \phi_o^3 \phi_\ell \rangle$$

$$\left. + 2\sum_{k,\ell,m} \lambda_k \lambda_\ell \lambda_m (E_k - E_o)(E_m - E_o) \langle \phi_o \phi_k \phi_\ell \phi_m \rangle + 0(\lambda^4) \right]. \tag{B8}$$

Expanding $\left(1 + \sum_k \lambda_k^2\right)^{-2}$ as a binomial series and multiplying out leaves

$$\text{Var}(\hat{\tilde{E}}_{MC}) = n^{-1}\left[\sum_{k,\ell} \lambda_k \lambda_\ell (E_k - E_o)(E_\ell - E_o)<\phi_o^2 \phi_k \phi_\ell> + 0(\lambda^3)\right]$$

$$+ \, 0(n^{-2}) \quad , \tag{4.13}$$

where only second order terms have been written out explicitly.

If Importance sampling has been used, then the corresponding equations for (4.3) and (4.4) are

$$\tilde{H} = \psi * H \psi / \rho \tag{B9}$$

$$\tilde{S} = \psi * \psi / \rho \tag{B10}$$

where ρ is a normalized and everywhere non-negative probability density (3.3). The equations for E_{var} (4.2) and $\text{Var}(\hat{\tilde{E}}_{MC})$ (4.10) still hold but where $\tilde{H}$ (B9) and $\tilde{S}$ (B10) replace H (4.3) and S (4.4), respectively. Hence, using the expansion in smallness parameters,

$$<\tilde{H}> = E_o<\phi_o^2/\rho> + \sum_k \lambda_k (E_o + E_k)<\phi_o \phi_k/\rho> + \sum_{k,\ell} \lambda_k \lambda_\ell E_k <\phi_k \phi_\ell/\rho> \tag{B11}$$

$$<\tilde{S}> = <\phi_o^2/\rho> + 2\sum_k \lambda_k <\phi_o \phi_k/\rho> + \sum_{k,\ell} \lambda_k \lambda_\ell <\phi_k \phi_\ell/\rho> \quad , \tag{B12}$$

which reduce to

$$\langle \tilde{H} \rangle = \langle H \rangle = E_o + \sum_k \lambda_k^2 E_k \tag{B13}$$

$$\langle \tilde{S} \rangle = \langle S \rangle = 1 + \sum_k \lambda_k^2 \tag{B14}$$

Hence the use of Importance sampling does not affect the expectation values and (4.7) holds for E_{var}.

The $Var(\hat{E}_{MC})$ for importance sampling is, after finding $Var(\tilde{H})$, $Var(\tilde{S})$, $Cov(\tilde{H},\tilde{S})$ for use in an analogous expression to (4.10) is

$$Var(\hat{E}_{MC}) = n^{-1} \left[\sum_{k,\ell} \lambda_k \lambda_\ell (E_k - E_o)(E_\ell - E_o) \langle \phi_o^2 \phi_k \phi_\ell / \rho^2 \rangle + 0(\lambda^3) \right]$$

$$+ 0(n^{-2}) \quad . \tag{B15}$$

The difference between (B15) and (4.13) to second order is the existence of the ρ^2 term in the denominator of the expectation value. Importance sampling does alter $Var(\hat{E}_{MC})$, however it is of note that (B15) is second order in λ.

Appendix C. Derivation of (4.19)

Given a normalized trial wavefunction, ψ_t (4.16), the variational energy, using (4.16) in (4.3), is

$$E_{var} \equiv \langle H \rangle \qquad\qquad (C1)$$

$$= E_o + \sum_{k=1}^{\infty} \lambda_k^2 (E_k - E_o) + O(\lambda^4) \qquad , \qquad (4.7)$$

where a binomial expansion of $\langle S \rangle^{-1}$ has been used. The variance of the Monte Carlo estimate is

$$Var(\widehat{E}_{MC}) \equiv n^{-1} Var(H) = n^{-1}\left(\langle H^2 \rangle - \langle H \rangle^2\right) \qquad . \qquad (C2)$$

After substituting (4.3) and (4.16) and then expanding in terms of λ,

$$Var(\widehat{E}_{MC}) = n^{-1}\left[E_o^2\left(\langle \phi_o^4 \rangle - 1\right) + 2\sum_{k=1}^{\infty} \lambda_k E_o (E_o + E_k)\langle \phi_o^3 \phi_k \rangle + O(\lambda^2) \right] \qquad ,$$

$$(4.19)$$

where terms of second and higher order in λ are implicit.

Appendix D. Derivation of (5.8) and (5.11)

In the DMC scheme, the trial wavefunction, ψ_T, and the time dependent wavefunction $\phi(\underset{\sim}{R},t)$ are

$$\psi_T = \phi_o + \sum_{k=1}^{\infty} \lambda_k \phi_k \qquad , \qquad (4.1)$$

and

$$\Phi = \sum_{k=0}^{\infty} \delta_k \phi_k \qquad , \qquad (5.2)$$

which provide the time dependent distribution function $f(\underset{\sim}{R},t)$ (5.4).

From (4.1) and (5.2), the normalization factor for $f(\underset{\sim}{R},t)$ is

$$\langle f(\underset{\sim}{R},t) \rangle^{-1} = \left[\delta_o + \sum_{k=1}^{\infty} \lambda_k \delta_k \right]^{-1}$$

$$= \delta_o^{-1} \left[1 - \sum_{k=1}^{\infty} \lambda_k \delta_k \delta_o^{-1} + \sum_{k,\ell=1}^{\infty} \lambda_k \lambda_\ell \delta_k \delta_\ell \delta_o^{-2} - \ldots \right] , \qquad (D1)$$

assuming

$$1 > \sum_{k} \lambda_k \delta_k \delta_o^{-1} \qquad .$$

From (5.6) and (5.7), the expected value of E_L given the distribution $f(\underset{\sim}{R},t)$ is

$$\langle E_L \rangle_f \equiv \langle (\psi_T^{-1} H \psi_T) f(\underset{\sim}{R},t) \rangle \langle f(R,t) \rangle^{-1} \quad . \tag{D2}$$

From (4.1), (5.2) and (5.4), the numerator is given by

$$\left(\delta_o E_o + \sum_{k=1}^{\infty} \lambda_k \delta_k E_k \right) \quad . \tag{D3}$$

Multiplying by the normalization factor, (D1), yields

$$\langle E_L \rangle_f = E_o + \sum_k \lambda_k \delta_k \delta_o^{-1} (E_k - E_o) + 0(\lambda^2 \delta^2 \delta_o^{-2}) \quad . \tag{5.8}$$

Since at infinite time $\delta_o = 1$, and $\delta_k = 0$ for all k,

$$\langle E_L \rangle_{f(\underset{\sim}{R},\infty)} = E_o \quad . \tag{D4}$$

By definition, the variance of E_L is

$$\text{Var}(E_L)_f = \langle E_L^2 \rangle_f - \langle E_L \rangle_f^2 \quad . \tag{D5}$$

where $\langle E_L^2 \rangle_f$ is defined by (5.9). Now

$$\langle E_L^2 \rangle_f \equiv \langle (\psi_T^{-1} H \psi_T)^2 f(\underset{\sim}{R},t) \rangle \langle f(\underset{\sim}{R},t) \rangle^{-1} \quad . \tag{D6}$$

96

From (4.1), (5.2) and (5.4), the numerator is given by

$$\delta_o E_o^2 + \sum_k \lambda_k \delta_k E_o (2E_k - E_o) + \sum_k \lambda_k^2 \delta_o (E_k - E_o)^2$$

$$+ \sum_{k,\ell,m} \lambda_k \lambda_\ell \Big(\lambda_m \delta_o (2E_o E_m - E_o^2 - E_k E_\ell)$$

$$+ \delta_m (E_k - E_o)(E_\ell - E_o) \Big) \langle \phi_k \phi_\ell \phi_m \phi_o^{-1} \rangle + \ldots \tag{D7}$$

Multiplying by the normalization factor (D1) and expanding yields

$$\langle E_L^2 \rangle_f = E_o^2 + 2 \sum_k \lambda_k \delta_k \delta_o^{-1} E_o (E_k - E_o)$$

$$+ \sum_k \lambda_k^2 (E_k - E_o)^2 + 0(\lambda^2 \delta_m \delta_o^{-1}) + 0(\lambda^3) \quad . \tag{D8}$$

Substitution (5.8) and (D8) into (D5), zero and first-order terms in λ cancel, leaving

$$\text{Var}(E_L)_f = \sum_{k=1}^{\infty} \lambda_k^2 (E_k - E_o)^2 + 0(\lambda^2 \delta_m \delta_o^{-1}) + 0(\lambda^3) \quad . \tag{5.11}$$

At infinite time, terms in δ disappear:

$$\text{Var}(E_L)_{f(R,\infty)} = \sum_{k=1}^{\infty} \lambda_k^2 (E_k - E_o)^2 + 0(\lambda^3) \quad . \tag{D9}$$

Appendix E. Importance Sampling for Hylleraas Coordinates [38]

The Importance sampling density function is obtained from a Slater-type orbital (STO)

$$\rho(s,t,u) = N\exp(-2\alpha s) \qquad , \qquad \text{(E1)}$$

where α is a constant and N is the normalization constant

$$N^{-1} = 2\pi^2 \int_0^\infty ds \int_0^s du \int_0^u dt \, u(s^2 - t^2)\exp(-2\alpha s) \qquad . \qquad \text{(E2)}$$

and

$$g = 2\pi^2 u(s^2 - t^2) \qquad \text{(E3)}$$

the Jacobian of transformation from Cartesian coordinates.

In order to select a set of random numbers from ρ (E1),

(i)   three random numbers p_{s_o}, p_{t_o}, p_{u_o} are chosen from a standard uniform probability density; these represent the probabilities of finding the two electrons at a point in configuration space with coordinates s_o, t_o and u_o, respectively.

(ii)   the ρ (E1) is integrated over t and u in order to obtain $\rho_1(s)$ and s_o is found such that

$$P_{s_o} = \int_0^{s_o} ds\rho_1(s) \qquad 0 \leqq s_o < \infty \qquad . \qquad \text{(E4)}$$

(iii) $\Gamma_2(u|s_o)$ is obtained by setting $s = s_o$ in (E1) and integrating over t; u_o is found such that

$$P_{u_o} = \int_0^{u_o} du\Gamma_2(u|s_o) \qquad 0 \leqq u_o \leqq s_o \qquad \text{(E5)}$$

(iv) $\Gamma_3(t|u_o,s_o)$ is obtained by setting $s = s_o$, $u = u_o$ in (E1) and t_o is found

$$P_{t_o} = \int_0^{t_o} dt\Gamma_3(t|s_o,u_o) \qquad 0 \leqq t_o \leqq u_o \qquad . \qquad \text{(E6)}$$

The $\rho_1(s)$ is expressed as a χ^2-function, so that s_o is computed from an inverse χ^2-function using IMSL [73] subroutine MDCHI. Densities $\Gamma_2(u|s_o)$ and $\Gamma_3(t|s_o,u_o)$ give rise to polynomial expressions in u_o and t_o, respectively; the largest positive root of (E5) and (E6) is solved for by Newton's method.

The Monte Carlo estimate of the quotient using Importance sampling is then, from (4.8)

$$\hat{E}_{MC} \cong n^{-1}\sum_{i=1}^{n} \tilde{H}(R^{(i)}) \Big/ n^{-1}\sum_{i=1}^{n} \tilde{S}(R^{(i)}) \qquad \text{(E7)}$$

where

$$\tilde{H} = g\psi^*H\psi/\rho \tag{E8}$$

$$\tilde{S} = g\psi^*\psi/\rho$$

where g is given by (E3).

The corresponding estimate of the variance of $\hat{E}_{MC}$ is from (4.10)

$$\hat{Var}(\hat{E}_{MC}) \cong (n\bar{\tilde{S}}^2)^{-1}\left(\hat{Var}(\tilde{H}) - 2\hat{E}_{MC}\hat{Cov}(\tilde{H},\tilde{S}) + \hat{E}_{MC}^2\hat{Var}(\tilde{S})\right) \quad . \tag{E10}$$

Appendix F. Precision of the Variance Estimate

It is apparent that the variance of an estimate can itself only be estimated, and so if we are interested in comparing the relative variances of estimators, then it would be appropriate to examine the precisions of these variance estimates. While one particular estimator may be known to have a smaller variance than another (depending on the sampling method chosen), the variance reduction achieved will be of little use if the error of the variance estimate is large.

Consider the ratio estimate $\hat{E}_{MC}$ (4.8). A sample estimate of the variance of $\widehat{Var}(\hat{E}_{MC})$ is approximately given by [36]:

$$\widehat{Var}\left(\widehat{Var}(\hat{E}_{MC})\right) \cong n^{-1}\left\{\hat{\beta}_z - 1 + 4\hat{V}_S^2 - 4\hat{\rho}_{z^2 H}(\hat{\beta}_z - 1)^{\frac{1}{2}}\hat{V}_S\right\}\widehat{Var}(\hat{E}_{MC}) \quad (F1)$$

where

$$z_i = H\left(R^{(i)}\right) - \hat{E}_{MC}S\left(R^{(i)}\right) \qquad (F2)$$

$$\hat{\beta}_z = n\sum_i z_i^4 \Big/ \left(\sum z_i^2\right)^2 \qquad (F3)$$

$$\hat{V} = \widehat{Var}(S)/\bar{S}^2 \qquad (F4)$$

and $\hat{\rho}_{z^2 H}$ is the estimated coefficient of correlation between z_i^2 and $H\left(R^{(i)}\right)$:

$$\hat{\rho}_{z^2 H} = n^{-1} \sum (z_i^2 - \overline{z^2}) \left[H(R^{(i)}) - \overline{H} \right] \Big/ \left[\hat{Var}^{\frac{1}{2}}(z^2) \hat{Var}^{\frac{1}{2}}(H) \right] \qquad . \quad (F5)$$

The square root of (F1) is used to create one standard deviation error bars on the estimated variance of $\hat{E}_{MC}$.

Appendix G.

Derivation of (8.19)

From the definition of H_{HF}^{TOT} (8.13)

$$H_{HF}^{TOT} = \sum_{i=1}^{N} H_i + \sum_{j \neq i} \sum_{i=1}^{N} (J_{ij} - K_{ij}) \quad . \tag{G1}$$

Using $\Phi(\underline{x})$ (8.15) to evaluate the expectation value

$$\langle \Phi(\underline{x}) | H_{HF}^{TOT} | \Phi(\underline{x}) \rangle = \sum_{i=1}^{N} \langle \Phi(\underline{x}) | H_i | \Phi(\underline{x}) \rangle$$

$$+ \sum_{j \neq i} \sum_{i=1}^{N} \langle \Phi(\underline{x}) | J_{ij} - K_{ij} | \Phi(\underline{x}) \rangle \quad . \tag{G2}$$

Appendix C in Blinder [56] shows that

$$\sum_{i=1}^{N} \langle \Phi(\underline{x}) | H_i | \Phi(\underline{x}) \rangle = \sum_{i=1}^{N} \langle \phi_{HP} | H_i | \sum_{r=1}^{N!} (-1)^r P_r \phi_{HP} \rangle$$

$$= \sum_{i=1}^{N} \langle \phi_i | H_i | \phi_i \rangle \quad . \tag{G3}$$

The second step in (G3) comes about by the properties of A and the symmetry of the H_i operator with respect to the electron coordinates.

The operators J_{ij} (8.8) and K_{ij} (8.9) do not affect the spin coordinates and are symmetrical with respect to the spatial coordinates.

Hence,

$$\sum_{j \neq i} \sum_{i=1}^{N} \langle \Phi(\underline{x}) | J_{ij} - K_{ij} | \Phi(\underline{x}) \rangle$$

$$= \sum_{j \neq i} \sum_{i=1}^{N} \int \cdots \int dx_1 \ldots dx_N \phi_1^*(1) \ldots \phi_N^*(N) (J_{ij} - K_{ij})$$

$$\times \sum_{r=1}^{N!} (-1)^r P_r \phi_1(1) \ldots \phi_N(N) \qquad . \tag{G4}$$

Only the permutations involving electrons i and j will contribute non-zero integrals and so the above can be written as

$$\sum_{j \neq i} \sum_{i=1}^{N} \int \cdots \int dx_1 \ldots dx_N \phi_1^*(1) \ldots \phi_N^*(N) \left(\int dx' \phi_j^*(x') r_{ix}^{-1} \phi_j(x') \right.$$

$$- \int dx' \phi_j^*(x') r_{ix}^{-1} P_{ij} \phi_j(x') \Bigg) \phi_1(1) \ldots [\phi_i(i)\phi_j(j) - \phi_j(i)\phi_i(j)]$$

$$\ldots \phi_N(N) \qquad . \tag{G5}$$

Carrying out the operations and separating the integrals leads to

$$\sum_{j\neq i}\sum_{i=1}^{N}\left\{\left[\int dx_1\phi_1^*(1)\phi_1(1)\right]\cdots\left[\iint dx_i dx'\phi_i^*(i)\phi_j^*(x')r_{ix}^{-1}\phi_i(i)\phi_j(x')\right]\right.$$

$$\left[\int dx_j\phi_j^*(j)\phi_j(j)\right]\cdots\left[\int dx_N\phi_N^*(N)\phi_N(N)\right]$$

$$-\sum_{j\neq i}\sum_{i=1}^{N}\left\{\left[\int dx_1\phi_1^*(1)\phi_1(1)\right]\cdots\left[\iint dx_i dx'\phi_i^*(i)\phi_j^*(x')r_{ix}^{-1}\phi_j(i)\phi_j(x')\right]\right.$$

$$\left[\int dx_j\phi_j^*(j)\phi_i(j)\right]\cdots\left[\int dx_N\phi_N^*(N)\phi_N(N)\right]$$

$$-\sum_{j\neq i}\sum_{i=1}^{N}\left\{\left[\int dx_1\phi_1^*(1)\phi_1(1)\right]\cdots\left[\iint dx_i dx'\phi_i^*(i)\phi_j^*(x')r_{ix}^{-1}\phi_j(i)\phi_i(x')\right]\right.$$

$$\left[\int dx_j\phi_j^*(j)\phi_j(j)\right]\cdots\left[\int dx_N\phi_N^*(N)\phi_N(N)\right]$$

$$+\sum_{j\neq i}\sum_{i=1}^{N}\left\{\left[\int dx_1\phi_1^*(1)\phi_1(1)\right]\cdots\left[\iint dx_i dx'\phi_i^*(i)\phi_j^*(x')r_{ix}^{-1}\phi_j(i)\phi_j(x')\right]\right.$$

$$\left[\int dx_j\phi_j^*(j)\phi_i(j)\right]\cdots\left[\int dx_N\phi_N^*(N)\phi_N(N)\right] \quad . \tag{G6}$$

But ϕ_i and ϕ_j are orthonormal, so second and fourth terms disappear, leaving

$$\sum_{j \neq i} \sum_{i=1}^{N} \left(\iint dx_i dx' \phi_i^*(i) \phi_j^*(x') r_{ix}^{-1} \phi_i(i) \phi_j(x') \right.$$

$$\left. - \iint dx_i dx' \phi_i^*(i) \phi_j^*(x') r_{ix}^{-1} \phi_j(i) \phi_i(x') \right)$$

$$= \sum_{j \neq i} \sum_{i=1}^{N} (J_{ij} - K_{ij}) \qquad . \tag{G7}$$

Combining (G2), (G3) and (G7) gives

$$\langle \Phi(\underline{x}) | H_{HF}^{TOT} | \Phi(\underline{x}) \rangle = \sum_{i=1}^{N} \langle \phi_i | H_i | \phi_i \rangle + \sum_{j \neq i} \sum_{i=1}^{N} (J_{ij} - K_{ij}) \quad . \tag{8.19}$$

Appendix H. Derivation of (8.32) and (8.33)

Given the wavefunction $\psi(\underset{\sim}{x})$ (8.28) and H_{HF}^{TOT} (8.24), $\langle H_{HF}^{TOT}\rangle$ is

$$\langle H_{HF}^{TOT}\rangle = \langle (U\Phi)^{*}\Big| -\tfrac{1}{2}\sum_{i}\nabla_i^{2} - \sum_{i}\sum_{\alpha}Z_{\alpha}r_{i\alpha}^{-1} + V_{HF}\Big| U\Phi\rangle \qquad . \tag{8.31}$$

operating ∇_i^{2} on $\psi(\underset{\sim}{x})$ gives

$$\langle H_{HF}^{TOT}\rangle = \langle (U\Phi)^{*}\Big| -\tfrac{1}{2}\sum_{i}\left(U\nabla_i^{2}\Phi + 2\vec{\nabla}_i U\cdot\vec{\nabla}_i\Phi + \Phi\nabla_i^{2}U\right)$$

$$+ \left(-\sum_{i}\sum_{\alpha}Z_{\alpha}r_{i\alpha}^{-1} + V_{HF}\right)U\Phi\rangle \qquad . \tag{H1}$$

Collecting terms,

$$\langle H_{HF}^{TOT}\rangle = \langle (U\Phi)^{*}\Big| U H_{HF}^{TOT}\Phi\rangle + \langle (U\Phi)^{*}\Big| -\tfrac{1}{2}\sum_{i}(2\vec{\nabla}_i U\cdot\vec{\nabla}_i\Phi + \Phi\nabla_i^{2}U)\rangle . \tag{H2}$$

The first term is $\sum_{i}\varepsilon_i\langle\psi^{*}\psi\rangle$ (to see this, expand Φ and substitute the definition of H_{HF}^{TOT} (8.13)), hence (H2) leads to (8.32). The second term in (H2) is, given real functions $U(\underset{\sim}{r}_{ij})$ and $\Phi(\underset{\sim}{x})$,

$$-\tfrac{1}{2}\sum_{i}\langle 2\Phi U\vec{\nabla}_i U\cdot\vec{\nabla}_i\Phi + \Phi^{2}U\nabla_i^{2}U\rangle \qquad . \tag{H3}$$

Since $\vec{\nabla}_i\Phi^{2} = 2\Phi\vec{\nabla}_i\Phi$, and employing (8.29), (H3) becomes

$$-\frac{1}{2}\sum_i \langle e^{-u}\vec{\nabla}_i e^{-u}\cdot\vec{\nabla}_i \Phi^2 + \Phi^2 e^{-u}\nabla_i^2 e^{-u}\rangle$$

$$= -\frac{1}{2}\sum_i \langle -e^{-2u}\vec{\nabla}_i u\cdot\vec{\nabla}_i \Phi^2 + \Phi^2 e^{-2u}(\vec{\nabla}_i u\cdot\vec{\nabla}_i u - \nabla_i^2 u)\rangle$$

$$= -\frac{1}{2}\sum_i \langle \frac{1}{2}\vec{\nabla}_i e^{-2u}\cdot\vec{\nabla}_i \Phi^2 - \Phi^2 e^{-2u}\vec{\nabla}_i u\cdot\vec{\nabla}_i u + \frac{1}{2}\Phi^2 \nabla_i^2 e^{-2u}\rangle$$

$$= -\frac{1}{2}\sum_i \langle \frac{1}{2}\vec{\nabla}_i \left(\Phi^2 \vec{\nabla}_i e^{-2u}\right) - \Phi^2 e^{-2u}\vec{\nabla}_i u\cdot\vec{\nabla}_i u\rangle \quad . \tag{H4}$$

Applying the Divergence Theorem [74] to the first integral in (H4):

$$\sum_i \int_V \vec{\nabla}_i \left[\Phi^2 \vec{\nabla}_i e^{-2u}\right]d\underline{V} = \sum_i \int_{S\infty} \Phi^2 \vec{\nabla}_i e^{-2u}d\underline{S} \tag{H5}$$

where the integration on the left-hand side is over all of space and on the right-hand side is over the surface at infinity. But Φ^2 must disappear at the infinite surface if Φ is a well-behaved wavefunction. Similarly, e^{-u}, e^{-2u} and $\vec{\nabla}_i e^{-2u}$ are zero at the infinite boundaries. Hence (H5) is zero, only the second term in (H4) contributes and

$$\langle H_{HF}^{TOT}\rangle = \sum_i \varepsilon_i \langle \psi^* \psi\rangle + \frac{1}{2}\sum_i \langle \vec{\nabla}_i u\cdot\vec{\nabla}_i u\rangle \quad . \tag{8.33}$$

Appendix I

Evaluation of $\int \chi_A(r_x)\chi_B(r_x)r_{1x}^{-1}dr_x$

From Rüdenberg [61], the charge distribution $\chi_A(r_x)\chi_B(r_x)$ for $\chi_A = \chi_B = \chi_{1s}$ where the orbital exponents $\zeta_A = \zeta_B = \zeta$, is

$$\chi_{1s}^2 = R^{-3}(R\zeta)^3\omega(\xi,\eta)\exp(-\zeta R\xi) \tag{I1}$$

where $\omega(\xi,\eta)$ is a polynomial in (ξ,η). Also, the relation

$$\tfrac{1}{2}\pi\left[\tfrac{1}{2}\left(1 + \delta_{MO}\right)\right]^{\frac{1}{2}}(\xi^2 - \eta^2)\omega(\xi,\eta) = \sum_{n=0}^{2}\sum_{j=0}^{2}\omega_{nj}\xi^n\eta^j \tag{I2}$$

is needed; ω_{nj} are coefficients in Table I in Rüdenberg.

The Neumann expansion of r_{12}^{-1} is used:

$$r_{12}^{-1} = 4R^{-1}\sum_{\ell=0}^{\infty}\sum_{m=-\ell}^{\ell}(-1)^m(\ell - |m|)!\left[(\ell + |m|)!\right]^{-1}$$

$$\times\, P_\ell^{|m|}(\xi_<)Q_\ell^{|m|}(\xi_>)P_\ell^{|m|}(\eta_1)P_\ell^{|m|}(\eta_2)$$

$$\times\, \exp(-im\phi_1)\exp(-im\phi_2) \tag{I3}$$

where $\xi_<$ and $\xi_>$ are the lesser and greater of ξ_1 and ξ_2, respectively. The normalized Legendre functions P_ℓ^m are given by

$$P_\ell{}^m(x) = \left((2\ell + 1)(\ell - m)!/2(\ell + m)!\right)^{\frac{1}{2}} P_\ell{}^m(x) \quad . \tag{I4}$$

Using (I1), (I3) and the volume element

$$dr = (R^3/8)(\xi^2 - \eta^2)d\xi d\eta d\phi \tag{I5}$$

in

$$A = \int \chi_A(r_x)\chi_B(r_x)r_{1x}{}^{-1}dr_x \tag{I6}$$

and setting $M = |m|$,

$$A = (2R)^{-1}(R\zeta)^3 \int_1^\infty\!\!\int_{-1}^1\!\!\int_0^{2\pi} d\xi_x d\eta_x d\phi_x \omega(\xi_x, \eta_x)(\xi_x{}^2 - \eta_x{}^2)$$

$$\times \exp(-\zeta R\xi_x)\sum_{\ell=M}^\infty (-1)^M(\ell - M)!\left((\ell + M)!\right)^{-1} P_\ell{}^M(\xi_<)Q_\ell{}^M(\xi_>)$$

$$\times P_\ell{}^M(\eta_1)P_\ell{}^M(\eta_x)\exp(-iM\phi_1)\exp(-iM\phi_x) \tag{I7}$$

For this $\chi_{1s}{}^2$ distribution, $M = 0$, so (I2) has a factor of $\pi/2$. Employing (I2) in (I7) and integrating over ϕ_x gives

$$A = 2R^{-1}(R\zeta)^3\sum_{\ell=0}^\infty P_\ell(\eta_1)\sum_{n=0}^2 \sum_{j=0}^2 \omega_{nj}\int_{-1}^1 d\eta_x \eta_x^j P_\ell(\eta_x)$$

$$\times \int_1^\infty d\xi_x \xi_x^n \exp(-\zeta R\xi_x)P_\ell(\xi_<)Q(\xi_>) \quad . \tag{I8}$$

Substitute $B_j^{0\ell}(0)$ (9.13) for the integral over η_x. Rearrange the integral over ξ_x so that

$$A = 2R^{-1}(R\zeta)^3 \sum_{\ell=0}^{\infty} P_\ell(\eta_1) \sum_{n=0}^{2} \sum_{j=0}^{2} \omega_{nj} B_j^{0\ell}(0)$$

$$\times \left[Q_\ell(\xi_1) \int_1^{\xi_1} d\xi_x \xi_x^n \exp(-\zeta R \xi_x) P_\ell(\xi_x) \right.$$

$$\left. + P_\ell(\xi_1) \int_{\xi_1}^{\infty} d\xi_x \xi_x^n \exp(-\zeta R \xi_x) Q_\ell(\xi_x) \right] \qquad . \qquad (I9)$$

From the definitions of $K_{\ell n}(\xi_1,\alpha)$ (9.11) and $L_{\ell n}(\xi_1,\alpha)$ (9.12) and P_ℓ (I4), (I9) becomes

$$A = 2R^{-1}(R\zeta)^3 \sum_{\ell=0}^{\infty} \left((2\ell+1)/2\right)^{\frac{1}{2}} P_\ell(\eta_1) \sum_{n=0}^{2} \sum_{j=0}^{2} \omega_{nj} B_j^{0\ell}(0)$$

$$\times \left\{ Q_\ell(\xi_1) K_{\ell n}(\xi_1,\zeta R) + P_\ell(\xi_1) L_{\ell n}(\xi_1,\zeta R) \right\} \qquad (9.10)$$

Following the expansion of $B_j^{0\ell}(0)$ and values of ω_{nj} given by Rüdenberg, only $\ell = 0,2$ terms are non-zero in (9.10). Appendices B and D in Harris [75] provide recursion formulae for $P_{\ell n}{}^m$, $Q_{\ell n}{}^m$, $K_{\ell n}{}^m$ and $L_{\ell n}{}^m$. In this work, $P_{\ell n}{}^m$, $Q_{\ell n}{}^m$, $K_{\ell n}{}^m$ were found using these formulae; the $L_{\ell n}{}^m$ were found using the explicit formula for L_{00}^0, a 24-point Gauss-Laguerre quadrature for $L_{\ell 0}{}^1$ and recursion formulae for other $L_{\ell n}{}^m$.

Appendix J.

Terms Arising From (9.19)

(a) Short R, where (9.5) holds

Using (9.7) and (9.8) as definitions of J_i, and the shortened

notations

$$[1234] = \phi_1(1)\phi_2(2)\phi_3(3)\phi_4(4) \tag{J1a}$$

$$[1432] = \phi_1(1)\phi_4(2)\phi_3(3)\phi_2(4) \tag{J1b}$$

$$[3214] = \phi_3(1)\phi_2(2)\phi_1(3)\phi_4(4) \tag{J1c}$$

$$[3412] = \phi_3(1)\phi_4(2)\phi_1(3)\phi_2(4) \tag{J1d}$$

where the (ordered) electrons, 1, ..., 4 are in orbitals designated by the

numbers in square brackets, then

$$\sum_{i=1}^{4} J_i \big([1234] - [1432] - [3214] + [3412]\big)$$

$$= [J_1(1) + J_2(2) + J_3(3) + J_4(4)]\big([1234] - [1432]\big)$$

$$+ [J_1(3) + J_2(4) + J_3(1) + J_4(2)]\big([3412] - [3214]\big) \qquad . \tag{J2}$$

Using similar notation to (J1), (9.15)-(9.18) for $K_i\phi_i$ and the

symbol

$$\kappa(1) = c_1 c_2 \left(\int \chi_A^2(r_x) r_{1x}^{-1} dr_x - \int \chi_B^2(r_x) r_{1x}^{-1} dr_x \right) \tag{9.15}$$

Appendix K. Comparison of Two Ratios

(a) Same Denominator

Given two ratios $\hat{R}$ and $\hat{R}'$ with the same denominator such that

$$\hat{R} = n^{-1}\sum_{i=1}^{n} y_i \Bigg/ n^{-1}\sum_{i=1}^{n} x_i = \bar{y}/\bar{x} \tag{K1}$$

$$\hat{R}' = n^{-1}\sum_{i=1}^{n} y_i' \Bigg/ n^{-1}\sum_{i=1}^{n} x_i = \bar{y}'/\bar{x} \tag{K2}$$

then

$$\hat{R} + \hat{R}' = (\bar{y} + \bar{y}')/\bar{x} \tag{K3}$$

Define d_i such that

$$d_i = (y_i + y_i') \qquad , \tag{K4}$$

then the estimated variance of $\hat{R} + \hat{R}'$ as given by (6.88) in Cochran [35] is

$$\text{Var}(\hat{R} + \hat{R}') \cong \left(n(n-a)\bar{x}^2\right)^{-1}\sum_{i=1}^{n} [d_i - (\hat{R} + \hat{R}')x_i]^2 \quad . \tag{K5}$$

Upon expansion of (K5) using (K1), (K2) and (K4),

$$\text{Var}(\hat{R} + \hat{R}') \cong \left(n(n-1)\bar{x}^2\right)^{-1}\sum_{i=1}^{n}\left[y_i^2 + 2y_i y_i' + y_i'^2 - 2\hat{R}y_i x_i\right.$$

$$- 2\hat{R}'y_i x_i - 2\hat{R}y_i'x_i - 2\hat{R}'y_i'x_i + 2\hat{R}\hat{R}'x_i^2$$

$$\left.+ \hat{R}^2 x_i^2 + \hat{R}'^2 x_i^2\right] \qquad . \qquad \text{(K6)}$$

Now, if $n \cong (n-1)$, then

$$\hat{\text{Var}}(y) \cong (n-1)^{-1}\sum_{i=1}^{n} y_i^2 - \bar{y}^2 \qquad \text{(K7a)}$$

$$\hat{\text{Var}}(y') \cong (n-1)^{-1}\sum_{i=1}^{n} y_i'^2 - \bar{y}'^2 \qquad \text{(K7b)}$$

$$\hat{\text{Var}}(x) \cong (n-1)^{-1}\sum_{i=1}^{n} x_i^2 - \bar{x}^2 \qquad \text{(K7c)}$$

$$\hat{\text{Cov}}(yy') \cong (n-1)^{-1}\sum_{i=1}^{n} y_i y_i' - \bar{y}\bar{y}' \qquad \text{(K7d)}$$

$$\hat{\text{Cov}}(yx) \cong (n-1)^{-1}\sum_{i=1}^{n} y_i x_i - \bar{y}\bar{x} \qquad \text{(K7e)}$$

$$\hat{\text{Cov}}(y'x) \cong (n-1)^{-1}\sum_{i=1}^{n} y_i'x_i - \bar{y}'\bar{x} \qquad \text{(K7f)}$$

Employing (K7) in (K6) gives

$$\hat{\text{Var}}(\hat{R} + \hat{R}') \cong \left(n\bar{x}^2\right)^{-1}\left[\hat{\text{Var}}(y) + \bar{y}^2 + 2\hat{\text{Cov}}(yy') + 2\bar{y}\bar{y}'\right.$$

$$+ \hat{\text{Var}}(y') + \bar{y}'^2 - 2\hat{R}\left\{\hat{\text{Cov}}(yx) + \bar{y}\bar{x}\right\}$$

$$- 2\hat{R}'\left\{\hat{\text{Cov}}(yx) + \bar{y}\bar{x}\right\} - 2\hat{R}\left\{\hat{\text{Cov}}(y'x) + \bar{y}'\bar{x}\right\}$$

$$\left. -2\hat{R}'\left\{\hat{\text{Cov}}(y'x) + \bar{y}'\bar{x}\right\} + \left\{\hat{R} + 2\hat{R}\hat{R}' + \hat{R}'\right\}\left(\hat{\text{Var}}(x) + \bar{x}^2\right)\right].$$

$$\text{(K8)}$$

But from (K1) and (K2),

$$\hat{R}\bar{y}\bar{x} = \bar{y}^2$$

$$\hat{R}'\bar{y}\bar{x} = \bar{y}\bar{y}' \ \ldots$$

$$\text{(K9)}$$

__etc.__, and so (K8) simplifies to

$$\hat{\text{Var}}(\hat{R} + \hat{R}') \cong \left(n\bar{x}^2\right)^{-1}\left[\hat{\text{Var}}(y) + 2\hat{\text{Cov}}(yy') + \hat{\text{Var}}(y')\right.$$

$$\left. - 2(\hat{R} + \hat{R}')\left\{\hat{\text{Cov}}(yx) + \hat{\text{Cov}}(y'x)\right\} + (\hat{R} + \hat{R}')^2\hat{\text{Var}}(x)\right]$$

$$\text{(K10)}$$

(b) Different Denominator

Given two ratios, $\hat{R}$ and $\hat{R}'$ with different denominators such that

$$\hat{R} = \bar{y}/\bar{x} \tag{K1}$$

$$\hat{R}' = \bar{y}'/\bar{x}' \tag{K11}$$

Then the variance of $(\hat{R} - \hat{R}')$ is

$$\text{Var}(\hat{R} - \hat{R}') = \text{Var}(\hat{R}) + \text{Var}(\hat{R}') - 2\text{Cov}(\hat{R}\hat{R}') \qquad . \tag{K12}$$

The variances of the ratios are estimated using the usual formula (4.10). The covariance term is, as in (6.90) in Cochran [35], estimated by

$$\hat{\text{Cov}}(\hat{R}\hat{R}') \cong (n\bar{x}\bar{x}')\hat{\text{Cov}}(y - \hat{R}x)\hat{\text{Cov}}(y' - \hat{R}'x')$$

$$\cong \left\{ n(n-1)\bar{x}\bar{x}' \right\} \sum_{i=1}^{n} \left\{ y_i y_i' - \hat{R}y_i'x_i - \hat{R}'y_i x_i' + \hat{R}\hat{R}'x_i x_i' \right\} .$$

$$\tag{K13}$$

Appendix L.   Derivation of $\vec{\nabla}_i u \cdot \vec{\nabla}_i u$ for (11.1) and (11.2)

(a) $u = \sum_{j>i} \sum_i a r_{ij}(1 + b r_{ij})^{-1}$ (L1)

This functional form of u is from the correlation function in (11.1). The gradient of the k-th coordinate of u is

$$\vec{\nabla}_k u = (\partial u / \partial x_k)\vec{i} + (\partial u / \partial y_k)\vec{j} + (\partial u / \partial z_k)\vec{k}$$ (L2)

where $\vec{i}$, $\vec{j}$, $\vec{k}$ denote unit vectors along the x, y, x axes, respectively. Consider $\partial u / \partial x_k$ using (L1):

$$\partial u / \partial x_k = \frac{\partial}{\partial x_k} \left[\sum_{j>i} \sum_i a r_{ij}(1 + b r_{ij})^{-1} \right]$$

$$= \sum_{j>k} \frac{\partial}{\partial x_k} \left[a r_{kj}(1 + b r_{kj})^{-1} \right]$$

$$= \sum_{j>k} a(1 + b r_{kj})^{-2} \partial r_{kj} / \partial x_k \quad .$$ (L3)

From the definition of r_{ij},

$$r_{ij} = [(x_i - x_j)^2 + (y_i - y_j)^2 + (z_i - z_j)^2]^{\frac{1}{2}}$$ (L4)

then

$$\partial r_{kj}/\partial x_k = (x_k - x_j)r_{kj}^{-1} \qquad .$$

(L5)

Substitution of (L5) into (L3) gives

$$\partial u/\partial x_k = a\sum_{j>k} (x_k - x_j)\left[r_{kj}(1 + br_{kj})^2\right]^{-1} \qquad .$$

(L6)

Analogous equations are found for $\partial u/\partial y_k$ and $\partial u/\partial z_k$.

Hence (L2) becomes

$$\vec{\nabla}_k u = a\sum_{j>k}\left[r_{kj}(1 + br_{kj})^2\right]^{-1}\vec{r}_{kj}$$

(L7)

and the dot product is

$$\vec{\nabla}_k u \cdot \vec{\nabla}_k u = \left[\sum_{j>k} a/(1 + br_{kj})^2\right]^2 \qquad .$$

(L8)

(b) $\quad u = \sum_{j>i} \sum_{i} \left[1 - \exp(-cr_{ij})\right] \bigg/ c^2 r_{ij}$ (L9)

This u is from (11.2). The $\vec{\nabla}_k u$ is given by (L2) where, for example,

$$\partial u/\partial x_k = \sum_{j>k} \frac{\partial}{\partial x_k} \left[1 - \exp(-cr_{kj})\right] \bigg/ c^2 r_{kj}$$

$$= \sum_{j>k} (x_k - x_j)\left[(cr_{kj} + 1)\exp(-cr_{kj}) - 1\right] \bigg/ c^2 r_{kj}{}^3 \qquad \text{(L10)}$$

using (L5). Then for (L9), following the steps in section (a),

$$\vec{\nabla}_k u \cdot \vec{\nabla}_k u = \left[\sum_{j>k} \left[(cr_{kj} + 1)\exp(-cr_{kj}) - 1\right] \bigg/ c^2 r_{kj}{}^2\right]^2 \quad . \qquad \text{(L11)}$$

Note that (L8) is composed of two-body components T_2

$$T_2 = \sum_{j>k} a^2/(1 + br_{kj})^4 \qquad \text{(L12)}$$

and three-body components T_3

$$T_3 = \sum_{j \neq \ell > k} a^2 \bigg/ (1 + br_{kj})^2 (1 + br_{k\ell})^2 \qquad . \qquad \text{(L13)}$$

Equation (L11) is similarly composed of T_2 and T_3 terms. The

calculation estimates

$$\hat{T}_2 \cong \langle \psi^* | T_2 | \psi \rangle / \langle \psi^* \psi \rangle \qquad \text{(L14)}$$

$$\hat{T}_3 \cong \langle \psi^* | T_3 | \psi \rangle / \langle \psi^* \psi \rangle \qquad \text{(L15)}$$

$$\hat{G} \cong \langle \psi^* | T_2 + T_3 | \psi \rangle / \langle \psi^* \psi \rangle$$

$$= \langle \vec{\nabla} u \cdot \vec{\nabla} u \rangle / \langle \psi^* \psi \rangle \qquad \text{(L16)}$$

where ψ is the correlated wavefunction.

Appendix M. Importance Sampling in Prolate Spheroidal Coordinates for $^1\Sigma_g^+\text{He}_2$

The one-electron sampling function for the He_2 molecule is taken to be a normalized sum of one-electron density functions centered on nucleus A and B :

$$\rho_{AB}(\xi,\eta,\phi) = N(R^3/8)(\xi^2 - \eta^2)\left(\rho_A(\xi,\eta) + \rho_B(\xi,\eta)\right) \tag{M1}$$

where

$$\rho_A(\xi,\eta) = \exp\left(-\zeta R(\xi + \eta)\right) \tag{M2a}$$

$$\rho_B(\xi,\eta) = \exp\left(-\zeta R(\xi - \eta)\right) \tag{M2b}$$

and

$$N^{-1} = (R^3/8)\int_0^{2\pi} d\phi \int_{-1}^1 d\eta \int_1^\infty d\xi(\xi^2 - \eta^2)(\rho_A + \rho_B) \quad . \tag{M3}$$

In order to select a set of random variables ξ_0,η_0,ϕ_0 from $\rho_{AB}(\xi,\eta,\phi)$, the following is carried out:

(i)   choose three random numbers $p(\xi_0)$, $p(\eta_0)$, $p(\phi_0)$ from a standard uniform probability density representing the probabilities of finding the electron at ξ_0, η_0 and ϕ_0, respectively.

(ii)   integrate ρ_{AB} (M1) over ϕ and η to obtain $\rho_1(\xi)$ and compute ξ_0 such that

$$p(\xi_0) = \int_1^{\xi_0} d\xi \rho_1(\xi) \quad , \qquad 1 \leqq \xi_0 < \infty \quad . \qquad (M4)$$

(iii)   set $\xi = \xi_0$ in ρ_{AB} (M1), integrate over ϕ to obtain $\Gamma(\eta|\xi_0)$ and solve for η_0:

$$p(\eta_0|\xi_0) = 2\int_0^{\eta_0} d\eta \Gamma(\eta|\xi_0), \qquad 0 \leqq \eta_0 \leqq 1 \quad , \qquad (M5)$$

(since $\Gamma(\eta|\xi_0)$ is symmetrical about $\eta = 0$). The range of η_0 is easily extended to $-1 \leqq \eta_0 \leqq 1$ by randomly setting $\eta_0 = -\eta_0$ with a 50% probability.

(iv)   since ϕ is independent of ξ and η, integrate over ξ and η to obtain $\rho(\phi)$ and solve for ϕ_0:

$$p(\phi_0) = \int_0^{\phi_0} d\phi \rho(\phi) = \phi_0/2\pi \qquad . \qquad (M6)$$

Note that $\rho_1(\xi)$ and $\Gamma(\eta|\xi)$ were found; it is equally valid to find $\rho_1(\eta)$ and $\Gamma(\xi|\eta)$ but there is no apparent significant difference between the difficulties in solving for η_0 conditional on ξ_0 and ξ_0 conditional on η_0.

Expressions $p(\xi_0)$ and $p(\eta_0|\xi_0)$ are complicated functions involving exponentials and polynomials in ξ_0 and η_0 respectively.:

$$p(\xi_0) = \tfrac{1}{4}\left\{\exp\left(-\zeta R(\xi_0 - 1)\right)[-\zeta R\xi_0{}^2 - 2\xi_0 + \zeta R - 2]\right.$$

$$\left. + 4 - \exp\left(-\zeta R(\xi_0 + 1)\right)[-\zeta R\xi_0{}^2 - 2\xi_0 + \zeta R + 2]\right\} \qquad \text{(M7)}$$

and

$$p(\eta_0|\xi_0) = \tfrac{1}{4}\left\{\exp(\zeta R\eta_0)[\exp(-\zeta R\xi_0)\left(\zeta^2 R^2(\xi_0{}^2 - \eta_0{}^2) + 2\zeta R\eta_0 - 2\right)]\right.$$

$$\left. - \exp(-\zeta R\eta_0)[\exp(-\zeta R\xi_0)\left(\zeta^2 R^2(\xi_0{}^2 - \eta_0{}^2) - 2\zeta R\eta_0 - 2\right)]\right\}$$

$$\text{(M8)}$$

Random numbers are generated using IMSL [73] subroutine GGUBFS. The roots of (M7) and (M8) are solved for using IMSL [73] subroutine ZSCNT.

A transformation of variables is used to constrain the ZSCNT search for η_0 in the range 0 to 1. It was found that at short internuclear separations, the sine^2 transformation,

$$y = \sin^2(y_t) \qquad \text{(M9)}$$

where y_t is the transformed variable, was suitable. At larger distances ($R > 9$ au) another transformation,

$$y = \exp(y_t)\left[\exp(-y_t) + \exp(y_t)\right]^{-1}$$

$$= \left(1 + \exp(-2y_t)\right)^{-1} \quad , \tag{M10}$$

had to be employed if ZSCNT were to successfully find η_0. At large R, the electron tends to spend a great deal of time close to either nucleus and so η_0 will be very close to unity, and the sine^2 transformation does not work well within the ZSCNT algorithm.

The coordinates of all four electrons were chosen independently using this method.

References

1. I. N. Levine, Quantum Chemistry (Allyn and Bacon, Boston, 1974), 2nd. ed., Chapt. 8.

2. I. M. Sobol', The Monte Carlo Method (The University of Chicago Press, Chicago and London, 1974).

3. J. M. Hammersley and D. C. Handscomb, Monte Carlo Methods (Methuen & Co., Ltd., London, 1965).

4. P. J. Davis and P. Rabinowitz, Methods of Numerical Integration (Academic Press, New York, 1975), Chapt. 5.

5. R. Y. Rubinstein, Simulation and the Monte Carlo Method (John Wiley & Sons, New York, 1981).

6. R. L. Gibbs and T. Dunn, J. Chem. Phys. 65 , 1192 (1976).

7. R. E. Lowther and R. L. Coldwell, Phys. Rev. A. 22, 14 (1980).

8. K. McDowell, Int. J. Quant. Chem.: Quant. Chem. Symp. 15, 177 (1981).

9. P. Hobza and R. Zahradník, Weak Intermolecular Interactions in Chemistry and Biology (Elsevier Scientific Publishing Co., Amsterdam, 1980).

10. R. L. Coldwell and R. E. Lowther, Int. J. Quant. Chem.: Quant. Chem. Symp. 12, 329 (1978).

11. D. C. Clary, Mol. Phys. 34, 793 (1977).

12. J. von Neumann and S. Ulam, Bulletin A. M. S., Abstract 51 - 9 -165 (1945).

13. N. Metropolis and S. Ulam, J. Amer. Stat. Assoc. 44, 335 (1949).

14. See, for example, references cited under: "Simulation" in Ref. 3; Chapt. 1 in Ref. 5.

15. See, for example, references cited under: "Integration" in Ref. 3; Chapt. 4 in Ref. 5.

16. See, for example, references cited under: "Nuclear Particle Transport" in Ref. 3.

17. W. W. Wood and J. J. Erpenbeck, Ann. Rev. Phys. Chem. 27, 319 (1976); G. M. Phillips, Comp. Phys. Comm. 20, 17 (1980).

18. R. Y. Rubenstein, op. cit. Chapt. 2.

19. Ibid. Chapt. 4.

20. H. Kahn in Symposium on Monte Carlo Methods ((H. A. Meyer, Ed.) John Wiley & Sons, New York, 1956), Chapt. 15; P. Heidelberger, IBM J. Res. Develop. 24, 570 (1980); P. Heidelberger in Current Issues in Computer Simulation (Academic Press, New York, 1979), Chapt. 17; S. S. Lavenberg, T. L. Moeller and C. H. Sauer, Oper. Res. 27, 134 (1979).

21. K. Binder in Monte Carlo Methods in Statistical Physics ((K. Binder, Ed.) Springer-Verlag, Berlin/New York, 1979).

22. D. M. Ceperley and M. H. Kalos in Monte Carlo Methods in Statistical Physics ((K. Binder, Ed.) Springer-Verlag, Berlin/New York, 1979).

23. J. P. Valleau and S. G. Whittington in Statistical Mechanics ((B. J. Berne, Ed.) Plenum Publishing Corp., 1977), Part A, Chapt. 4 and 5.

24. Accuracy is the difference between the sample estimate and the true value; precision is a measure of how closely the sample estimate can be reproduced. Although accuracy is the main interest in an experiment, precision is what is actually measured (unless the true value of the estimate is known).

25. P. J. Reynolds, D. M. Ceperley, B. J. Alder and W. A. Lester, J. Chem. Phys. 77, 5593 (1982).

26. J. W. Moskowitz, K. E. Schmidt, M. A. Lee and M. H. Kalos, J. Chem. Phys. 77, 349 (1982).

27. J. B. Anderson, J. Chem. Phys. _74_, 6307 (1981).

28. C. E. Clark, Oper. Res. _9_, 603 (1961).

29. R. Y. Rubinstein, _op. cit._ Chapt. 3.

30. T. Sasaki, SIAM J. Numer. Anal. _15_, 938 (1978).

31. S. M. Rothstein and H. L. Gordon, unpublished work.

32. J. Goodisman and W. Klemperer, J. Chem. Phys. _38_, 721 (1963).

33. M. Tin, J. Amer. Stat. Assoc. _60_, 294 (1965).

34. H. Cramér, _Mathematical Methods of Statistics_ (Princeton Univ. Press, Princeton, 1963), Chapt. 17.

35. W. A. Cochran, _Sampling Techniques_ (Wiley, New York, 1977), 3rd. ed., Chapt. 6.

36. M. H. Hansen, W. N. Hurwitz and W. G. Madow, _Sample Survey Methods and Theory_ (Wiley, New York, 1953), Vol. I and II, Chapt. 4.

37. H. A. Bethe and E. E. Salpeter, _Quantum Mechanics of One- and Two-Electron Atoms_ (Springer-Verlag, Berlin, 1957).

38. H. L. Gordon, S. M. Rothstein and T. R. Proctor, J. Comp. Phys. _47_, 375 (1982).

39. VA09A, Algorithm 60, Quantum Chemistry Program Exchange, Indiana University, Bloomington, Indiana 47401.

40. H. Ratajczak and W. J. Orville-Thomas in _Molecular Interactions_ ((H. Ratajczak and W. J. Orville-Thomas, Ed.) John Wiley & Sons, New York, 1980), Vol. 1, Chapt.1; K. Morokuma and K. Kitaura, _loc. cit._ Chapt. 2.

41. B. Jeziorski and W. Kołos in _Molecular Interactions_ ((H. Ratajczak and W. J. Orville-Thomas, Ed.) John Wiley & Sons, New York, 1982), Vol. 3, Chapt. 1.

42. W. Kutzelnigg, Faraday Discuss. Chem. Soc. _62_, 185 (1977); P. A. Christiansen, K. S. Pitzer _et. al._, J. Chem. Phys. _75_, 5410 (1981).

43. J. S. Winn, J. Chem. Phys. $\underline{74}$, 608 (1981).

44. G. Scoles, Ann. Rev. Phys. Chem. $\underline{31}$, 81 (1980).

45. A. D. Buckingham in Vibrational Spectroscopy of Molecular Liquids and Solids ((S. Bratos and R. M. Pick, Ed.) NATO Advanced Study Institute, 1980), Series B $\underline{56}$, pp. 1 - 41.

46. G. G. Maitland, M. Rigby, E. B. Smith and W. A. Wakeham, Inter-molecular Forces, Their Origin and Determination (Clarendon Press, Oxford, 1981).

47. P. Hobza and R. Zahradník, op. cit. Chapt. II.1.b.

48. S. F. Boys and F. Bernardi, Mol. Phys. $\underline{19}$, 553 (1970).

49. N. S. Ostlund and D. L. Merrifield, Chem. Phys. Lett. $\underline{39}$, 612 (1976).

50. D. R. McLaughlin and H. F. Schaefer III, Chem. Phys. Lett. $\underline{12}$, 244, (1971).

51. P. Bertoncini and A. C. Wahl, J. Chem. Phys. $\underline{58}$, 1259 (1973).

52. B. Liu and A. D. McLean, J. Chem. Phys. $\underline{59}$, 4557 (1973).

53. P. D. Dacre, Phys. Lett. $\underline{50}$, 147 (1977).

54. P. G. Burton, J. Chem. Phys. $\underline{70}$, 3112 (1979).

55. A. L. J. Burgmans, J. M. Farrer and Y. T. Lee, J. Chem. Phys. $\underline{64}$, 1345 (1976).

56. S. M. Blinder, Amer. J. Phys. $\underline{33}$, 431 (1965).

57. G. E. Brown, Many-Body Problems (North-Holland, Amsterdam/London, 1972), pp. 25 - 33.

58. G. Arfkin, Mathematical Methods for Physicists (Academic Press, New York/London, 1966).

59. The CDC version of HONDO was provided by H. F. King and N. Camp, Dept. of Chemistry, State University of New York , Buffalo. HONDO was modified by E. M. Bölük to run on the BURROUGHS 6700. Both IBM and CDC versions of HONDO are available: M. Dupuis, J. Rys

and H. F. King, HONDO 67, programs 338 and 336, Quantum Chemistry Program Exchange, University of Indiana, Bloomington, Indiana 47401.

60. C. C. J. Roothan, J. Chem. Phys. $\underline{19}$, 1445 (1951).

61. K. Rüdenberg, J. Chem. Phys. $\underline{19}$, 1459 (1951).

62. R. Y. Rubenstein, op. cit. pp. 124 - 126.

63. R. Jastrow, Phys. Rev. $\underline{98}$, 1479 (1955).

64. N. C. Handy, J. Chem. Phys. $\underline{51}$, 3205 (1969); S. F. Boys and N. C. Handy, Proc. R. Soc. (London) Series A $\underline{309}$, 209 (1969); $\underline{310}$, 43 (1969); $\underline{310}$, 63 (1969); $\underline{311}$, 309 (1969).

65. W. Kolos and C. C. J. Roothan, Rev. Mod. Phys. $\underline{32}$, 205 (1960).

66. M. Aubert-Frécon and C. Le Sech, J. Chem. Phys. $\underline{74}$, 2931 (1981); J. W. Moskowitz and M. H. Kalos, Int. J. Quant. Chem. $\underline{20}$, 1107 (1981).

67. M. S. Becker, A. A. Broyles and T. Dunn, Phys. Rev. $\underline{175}$, 224 (1968).

68. F. A. Stevens, Jr. and M. A. Pokrant, Phys. Rev. A $\underline{8}$, 990 (1973).

69. C. C. Spicer, Appl. Stat. $\underline{21}$, 226 (1972).

70. P. Bertoncini and A. C. Wahl, Phys. Rev. Lett. $\underline{25}$, 991 (1970).

71. D. L. Iglehart and P. A. W. Lewis, "Variance Reduction for Regenerative Simulations, I: Internal Control and Stratified Sampling for Queues," Technical Report No. 86 - 22, Control Analysis Corp., Palo Alto, Calif. (1976).

72. P. S. Laplace, Théorie Analytique des Probabilités (Paris, 1920), 3rd. ed.

73. International Mathematical and Statistical Libraries, Inc. IMSL, 1982, Version 67, Houston, Texas.

74. G. B. Thomas, Jr., Calculus and Analytic Geometry (Addison-Wesley Publishing Co., Reading, Mass., 1968), 4th. ed., Chapt. 17.6.

75. F. E. Harris and H. H. Michels, Adv. Chem. Phys. $\underline{13}$, 205 (1967).