

Guidingfunction optimization in biasedselection Monte Carlo calculations

S. A. Alexander, R. L. Coldwell, and John D. Morgan III

Citation: *The Journal of Chemical Physics* **97**, 8407 (1992); doi: 10.1063/1.463410

View online: <http://dx.doi.org/10.1063/1.463410>

View Table of Contents: <http://scitation.aip.org/content/aip/journal/jcp/97/11?ver=pdfcov>

Published by the [AIP Publishing](#)

Articles you may be interested in

[Biased Monte Carlo Methods](#)

AIP Conf. Proc. **690**, 99 (2003); 10.1063/1.1632121

[A biased Monte Carlo technique for calculation of the density of states of polymer films](#)

J. Chem. Phys. **116**, 7238 (2002); 10.1063/1.1463422

[Biased Monte Carlo optimization of protein sequences](#)

J. Chem. Phys. **113**, 2489 (2000); 10.1063/1.482067

[Quantum freeenergy calculations: Optimized Fourier pathintegral Monte Carlo computation of coupled vibrational partition functions](#)

J. Chem. Phys. **97**, 3647 (1992); 10.1063/1.462948

[Optimized trial functions for quantum Monte Carlo](#)

J. Chem. Phys. **92**, 597 (1990); 10.1063/1.458409



Guiding-function optimization in biased-selection Monte Carlo calculations

S. A. Alexander and R. L. Coldwell

Institute for Astrophysics and Planetary Exploration, Department of Physics, University of Florida, Gainesville, Florida 32611

John D. Morgan III

Institute for Theoretical Atomic and Molecular Physics, Harvard-Smithsonian Center for Astrophysics, 60 Garden Street, Cambridge, Massachusetts 02138 and Department of Physics and Astronomy, University of Delaware, Newark, Delaware 19716^{a)}

(Received 18 May 1992; accepted 17 August 1992)

We show that the guiding function, which is used to generate the Monte Carlo integration points in a biased-selection calculation, can be optimized so as to minimize the statistical error of an expectation value. For excited states we introduce an orthogonalizing part into the minimization functional to prevent ground and excited states from mixing.

I. INTRODUCTION

Variational Monte Carlo is a method of computing the expectation value of an operator (such as the Hamiltonian) with respect to a trial wave function, Ψ , using Monte Carlo integration.¹⁻¹⁴ Typically, the adjustable parameters in Ψ are optimized so as to minimize some functional (such as the variance). In Ref. 13 we showed that the choice of minimization functional effects the accuracy of the total energy. Not surprisingly, the form of the trial wave function also has a large influence on the total energy and other properties (see, for example, Ref. 14). In this paper we demonstrate that the procedure used to generate the Monte Carlo integration points can significantly effect the accuracy of a variational Monte Carlo calculation. Typically, these points are generated using a guiding (i.e., importance) function or from the Markov chain algorithm which effectively selects the Monte Carlo integration point r_i with probability $\Psi_i(r_i)$.^{2,8,10-12} Earlier calculations using the biased-selection method chose this guiding function on the basis of general physical principles.^{2-7,9,13,14} In Sec. II we show that the form of the guiding function, like the wave function, can be adjusted so as to minimize the variance of the total energy. In general, the guiding function places configurations where a combination of $\Psi_i(r_i)^2$ and its error is large rather than strictly according to $\Psi_i(r_i)^2$. Section III describes the results of this procedure for the helium 1^1S and 2^3S states as well as the H_2 and H_3^+ ground states. In Sec. IV we extend our procedure to examine the helium 2^1S excited state and the H_2 first singlet excited state. Unlike our earlier calculations,¹³ we now require that these excited states be explicitly orthogonalized to all lower states during the optimization step. We discuss in some detail the advantages and disadvantages of this procedure. In Sec. V we show that guiding functions can also be optimized so as to minimize the error in properties other than the total energy.

II. GUIDING-FUNCTION OPTIMIZATION

As previously described in Ref. 13, biased-selection Monte Carlo uses a small set of configurations, \mathbf{x}_i , and weights, $w(\mathbf{x}_i)$, to evaluate multidimensional integrals,

$$I = \frac{1}{N} \sum_{i=1}^N f(\mathbf{x}_i) / w(\mathbf{x}_i) \rightarrow \int f(\mathbf{x}) d\mathbf{x} \quad N \rightarrow \infty \quad (1)$$

Each configuration is a collection of electronic coordinates, e.g., $\mathbf{x}_i = (r_1, r_2)$ for a two-electron system, and contains the usual Cartesian, $\mathbf{r}_j = (x_j, y_j, z_j)$, or spherical, $\mathbf{r}_j = (r_j, \theta_j, \phi_j)$, components. The weight function is the relative probability of choosing a given configuration and exactly compensates for not having chosen it randomly.

For a two-electron atom (and probably for any system with up to ten electrons) it is sufficient to take r_1 and r_2 "biased as random." This means finding a point (r_j, θ_j, ϕ_j) such that

$$R = \int_0^{r_j} g(s) ds, \quad \Theta = \theta_j, \quad \Phi = \phi_j, \quad (2)$$

where R , Θ , and Φ are random numbers chosen from the range

$$R = \left[0, \int_0^\infty g(s) ds \right], \quad \Theta = [0, 2\pi] \quad \text{and} \quad \Phi = [-\pi, \pi], \quad (3)$$

and $g(s)$ is some predetermined guiding function. Since the coordinates of each electron are chosen independently, the weight of each configuration is given by

$$w(\mathbf{x}_i) = \hat{w}(\mathbf{x}_1) \hat{w}(\mathbf{x}_2), \quad (4)$$

where

$$\hat{w}(\mathbf{x}_j) = g(r_j) / \left[4\pi r_j^2 \int_0^\infty g(s) ds \right]. \quad (5)$$

The factor $4\pi r_j^2$ is needed here to compensate for choosing the points in spherical coordinates in Eq. (2).

In its simplest form, the introduction of a guiding function is equivalent to importance sampling.¹⁵ If $g(\mathbf{x})$ is large where $f(\mathbf{x})$ is large and small where $f(\mathbf{x})$ is small, then the configurations will tend to be chosen from those regions which contribute most to the integral. The only restriction

^{a)}Permanent address.

on the choice of $g(x)$ is that it be positive definite everywhere and computable. This choice does not effect the final result, but for a given wave function it can have a dramatic effect on the number of configurations needed to accurately estimate the expectation value and its error. The maximum improvement will occur when the form of the guiding function is optimized so as to minimize the statistical error. Such a calculation is straightforward:

(1) Pick an initial trial wave-function form.

(2) Pick an initial form for $g(r)$. Our general form, a linear spline (a set of points which are connected by a line) can approximate almost any function. The height and position of these points completely describes the guiding function.

(3) Using the current guiding function, generate a small set of configurations. Optimize the parameters in the current trial wave function so as to minimize the variance functional^{2,13}

$$\sigma^2 = \sum_i [(H\Psi_i - E_{in}\Psi_i)^2 \Psi_i^2 / w_i^2] / \left[\sum_i (\Psi_i^2 / w_i) \right]^2, \quad (6)$$

where $\Psi_i = \Psi_i(\mathbf{x}_i)$, $w_i = w(\mathbf{x}_i)$, and E_{in} is a constant which is fixed at a value close to the desired state in order to start the optimization in the proper region. The energy produced by this wave function is given by

$$\langle H \rangle = \sum_i (\Psi_i H \Psi_i / w_i) / \sum_i (\Psi_i^2 / w_i). \quad (7)$$

Note that both the variance, Eq. (6), and the energy, Eq. (7), are computed with the current set of configurations. If this energy and variance agree with those from the previous iteration (within some tolerance), then quit.

(4) Adjust the height, position and number of points describing the guiding function so as to minimize the functional (see the Appendix for a derivation),

$$\frac{\sum_i \frac{[H\Psi_i(\mathbf{x}_i) - E_{in}\Psi_i(\mathbf{x}_i)]^2 \Psi_i(\mathbf{x}_i)^2}{w(\mathbf{x}_i)} \frac{\Psi_i(\mathbf{y}_i)}{\omega(\mathbf{y}_i)}}{\left[\sum_i \frac{\Psi_i(\mathbf{x}_i)^2}{w(\mathbf{x}_i)} \right]^2} \quad (8)$$

Here \mathbf{x}_i and $w(\mathbf{x}_i)$ are generated [Eqs. (2)–(5)] with the guiding function from step 3 using a fixed set of random numbers. The new configurations \mathbf{y}_i and $\omega(\mathbf{y}_i)$ are generated with the new guiding function and the same set of random numbers. The trial wave function is the one computed in step 3 and is not altered here.

(5) Return to step 3.

It is easy to find guiding functions such that Eqs. (6) and (7) are dominated by a single configuration (and thus yield meaningless results). For this reason, as emphasized in Ref. 13, it is useful to calculate the effective number of configurations which contribute to the normalization integral

$$N_{\text{eff}} = \left[\sum_i (\Psi_i^2 / w_i) \right]^2 / \sum_i (\Psi_i^4 / w_i^2). \quad (9)$$

This quantity will be roughly 1.0 when only a few of the points dominate the integral and it will be roughly N when

TABLE I. The energy of the helium 1 1S state (in hartrees). N is the number of configurations.

N	Ref. 13 ^a	Optimized $g(r)$
16 000	−2.903 734 1 ± 0.000 095 2	−2.903 680 4 ± 0.000 023 2
64 000	−2.903 750 1 ± 0.000 050 1	−2.903 711 8 ± 0.000 011 8
256 000	−2.903 726 5 ± 0.000 023 4	−2.903 713 1 ± 0.000 005 9
1 024 000	−2.903 707 4 ± 0.000 011 2	−2.903 716 8 ± 0.000 002 9
Pekeris (Ref. 17)		−2.903 724 375 ^b
Rybak, Szalewicz, and Jezierski (Ref. 18)		−2.903 723 8 ^c
Thakkar and Smith (Ref. 19)		−2.903 724 363 ^d
Baker <i>et al.</i> (Ref. 20)		−2.903 724 377 034 118 4 ^e
Umrigar, Wilson, and Wilkins (Ref. 8)		−2.903 722(2) ^f
Harrison and Handy (Ref. 21)		−2.903 74(5) ^g

^aVariance functional minimized using $g(r) = 23.404r^2e^{-3r}$ when $r = [0, 1]$ and $g(r) = \text{maximum}(23.404r^2e^{-3r}, 0.042r^4e^{-r})$ when $r = [1, 100]$.

^b1078 functions in perimetric coordinates.

^c100 fully optimized Gaussian-type geminals.

^d66 random-tempered Slater-type geminals.

^e476 Hylleraas functions with logarithmic terms.

^fVariational Monte Carlo.

^gDiffusion Monte Carlo.

all the points make roughly equal contributions. In previous calculations we found that N_{eff} had to be on the order of 10 times the number of adjustable parameters and between $0.2N$ and $0.4N$ to perform an accurate optimization. If N_{eff} is too small then too many points are being wasted sampling unimportant regions instead of lowering σ . If N_{eff} is too large then not enough points are sampling the other regions to see if any of them are important. One test of our guiding-function optimization process will be to see how well our algorithm automatically reproduces these rules of thumb.

III. GROUND-STATE RESULTS

In order to most easily show that optimizing the guiding function can effect the total energy and its variance, we examine the same systems as in Ref. 13. The present calculations were performed in exactly the same way as these earlier calculations except that we now optimize both the trial wave function and the guiding function so as to minimize the variance. For the helium 1 1S state our trial wave function is

$$\Psi_t = (1 + P_{12}) \sum_{k=1}^{34} a_k [e^{-\alpha r_1 - \beta r_2} r_{12}^{b_k} (r_1 - r_2)^{c_k} (r_1 + r_2)^{d_k}], \quad (10)$$

where P_{12} is the operator which interchanges the coordinates of the two electrons; α , β , and a_k are the parameters to be optimized and b_k , c_k , and d_k are predetermined integer constants. Table I shows that optimizing the guiding function reduces the statistical error by almost a factor of 3 and also improves the expectation value. During the optimization process the value of N_{eff} rose from 43% for the original guiding function to 49% for the optimized one. A plot of both functions is given in Fig. 1. The major differences between these two functions are at small r and large r . The reason for this is that an ideal guiding function should place a large number of points where Ψ_t is signifi-

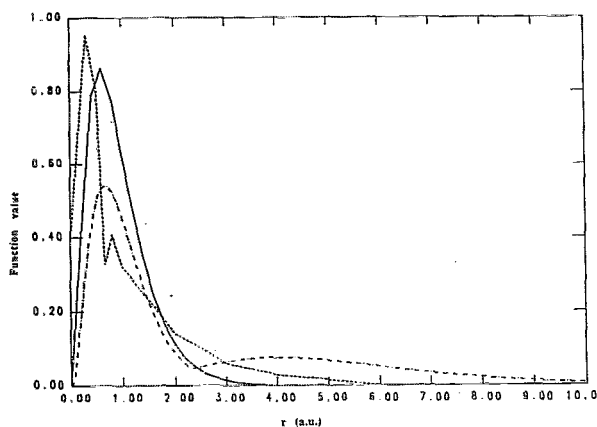


FIG. 1. Plot of the optimized guiding function (---), the original guiding function (---), and the radial density of the wave function, Eq. (11) (solid line), for the He $1S$ state. The area of each curve has been normalized to one.

cant and where $H\Psi_i/\Psi_i$ differs most from E . In those regions where the wave function is good, few points are needed. Optimizing the guiding function places more configurations into those regions where the variance is large, thereby decreasing the importance of each point quadratically—because the variance goes as Ψ^4/w^2 while the number of points increases as Ψ^2/w , i.e., linearly (this is also why variance optimization usually gives a good estimate of the energy). The reason that the optimized guiding function is sizable at large r is because variance minimization generally tends to push the error in evaluating $H\Psi_i - E\Psi_i$ into regions where the wave function is small. Because the helium potential is singular at the origin, because the wave function of this system has its maximum at the origin, and because the form of our trial wave function does not explicitly satisfy the cusp condition at the origin, the optimized guiding function wants to put a large number of configurations there. We have examined several wave-function forms which *do* explicitly satisfy the cusp condition at the origin and found that for a fixed number of constants they raised the total energy and/or the variance.

Also shown in Fig. 1 is a plot of the probability of finding an electron at a distance r from nucleus A . This quantity was calculated using the expression

$$P(r) = \int \Psi_i(\mathbf{r}_1, \mathbf{r}_2)^2 \delta(|\mathbf{r}_{1A}| = r) d^3r_1 d^3r_2. \quad (11)$$

Figure 1 shows that this function only superficially resembles the optimal guiding function.

For our calculations on the helium triplet ground state, we used the trial wave function,

$$\Psi_t = (1 - P_{12}) \sum_{k=1}^{34} a_k [e^{-\alpha r_1 - \beta r_2} r_{12}^{b_k} (r_1 - r_2)^{c_k} (r_1 + r_2)^{d_k}]. \quad (12)$$

During the optimization process the value of N_{eff} dropped from 39% for the previous guiding function to 35% for the

TABLE II. The energy of the helium 2^3S state (in hartrees). N is the number of configurations.

N	Ref. 13 ^a	Optimized $g(r)$
16 000	$-2.175\,213\,9 \pm 0.000\,027\,5$	$-2.175\,219\,4 \pm 0.000\,018\,4$
64 000	$-2.175\,218\,2 \pm 0.000\,014\,1$	$-2.175\,219\,5 \pm 0.000\,009\,2$
256 000	$-2.175\,202\,0 \pm 0.000\,006\,8$	$-2.175\,219\,3 \pm 0.000\,004\,6$
1 024 000	$-2.175\,207\,6 \pm 0.000\,003\,3$	$-2.175\,204\,1 \pm 0.000\,002\,3$
Pekeris (Ref. 17)		$-2.175\,229\,378\,237^b$
Thakkar and Smith (Ref. 22)		$-2.175\,229\,376^c$
Drake (Ref. 23)		$-2.175\,229\,378\,236\,790\,7^d$
Umrigar, Wilson, and Wilkins (Ref. 8)		$-2.175\,226(2)^e$
Zhang and Kalos (Ref. 24)		$-2.175\,243(66)^f$

^aVariance functional minimized using $g(r) = 23.404r^2e^{-3r}$ when $r = [0, 1]$ and $g(r) = \text{maximum}(23.404r^2e^{-3r}, 0.042r^4e^{-r})$ when $r = [1, 100]$.

^b1078 functions in perimetric coordinates.

^c55 random-tempered Slater-type geminals.

^d616 Hylleraas functions with multiple sets of exponents.

^eVariational Monte Carlo.

^fDiffusion Monte Carlo.

optimized function. The values in Table II show, however, that optimizing the guiding function improves the variance only slightly. A plot of both functions is given in Fig. 2 together with $P(r)$. Like the singlet ground state, the optimized guiding function differs from the original one mostly at small r and very large r .

For the H_2 ground state at internuclear distance 1.4 a.u. we took our trial wave function to have the form

$$\Psi_t = (1 + P_{12}) \sum_{k=1}^{126} a_k [e^{-\alpha r_{1A} - \beta r_{2B}} r_{12}^{b_k} (r_{1A} - r_{1B})^{c_k} (r_{1A} + r_{1B})^{d_k} (r_{2A} - r_{2B})^{e_k} (r_{2A} + r_{2B})^{f_k}], \quad (13)$$

where A and B are the positions of the two nuclei. As pointed out in Ref. 13, for molecular systems the function used to generate the weights [Eq. (5)] is replaced by the expression

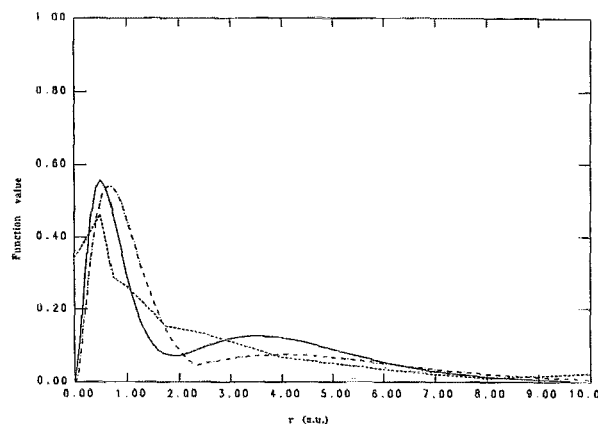


FIG. 2. Plot of the optimized guiding function (---), the original guiding function (---), and the radial density of the wave function, Eq. (11) (solid line), for the He $2S$ state. The area of each curve has been normalized to one.

TABLE III. The energy of the H_2 ground state (in hartrees). N is the number of configurations. The internuclear distance is 1.4 a.u.

N	Ref. 13 ^a	Optimized $g(r)$
16 000	-1.174 540 6 ± 0.000 100 1	-1.174 442 6 ± 0.000 067 9
64 000	-1.174 399 8 ± 0.000 050 1	-1.174 427 9 ± 0.000 034 0
256 000	-1.174 445 2 ± 0.000 025 4	-1.174 453 8 ± 0.000 017 2
Kolos, Szalewicz, and Monkhorst (Ref. 25)		-1.174 475 668 ^b
Frye, Lie, and Clementi (Ref. 26)		-1.174 474 67 ^c
Alexander <i>et al.</i> (Ref. 27)		-1.174 474 85 ^d
Sun <i>et al.</i> (Ref. 11)		-1.172 2(1) ^e
Anderson (Ref. 28)		-1.174 517(47) ^f

^aVariance functional minimized using $g(r) = 12.371r^2e^{-2r}$ when $r = [0, 1]$ and $g(r) = \text{maximum}(12.371r^2e^{-2r}, 0.042r^4e^{-r})$ when $r = [1, 100]$.

^b249 explicitly correlated functions in elliptical coordinates.

^cLinear Gaussian-type geminal calculation.

^d900 random-tempered Gaussian-type geminals.

^eVariational Monte Carlo.

^fDiffusion Monte Carlo.

$$\hat{w}(\mathbf{x}_j) = \frac{1}{\# \text{ atoms}} \sum_{k=1}^{\# \text{ atoms}} g(|\mathbf{r}_j - \mathbf{R}_k|) [4\pi |\mathbf{r}_j - \mathbf{R}_k|^2 \times \int_0^\infty g(s) ds]^{-1}. \quad (14)$$

In Table III we see that the statistical error improves by almost 33% and that the expectation value also improves. During the optimization process the value of N_{eff} dropped from 75% for the original guiding function to 71% for the optimized function. A plot of both functions is given in Fig. 3 together with the probability of finding an electron at a distance r from nucleus A . The peak around 0.7 a.u. suggests that the optimized guiding function is trying to place a number of configurations in the center of the molecule.

At its equilibrium configuration the ground state of the H_3^+ ion has D_{3h} symmetry and a bond distance of 1.65 a.u. For this reason our trial wave function has the form

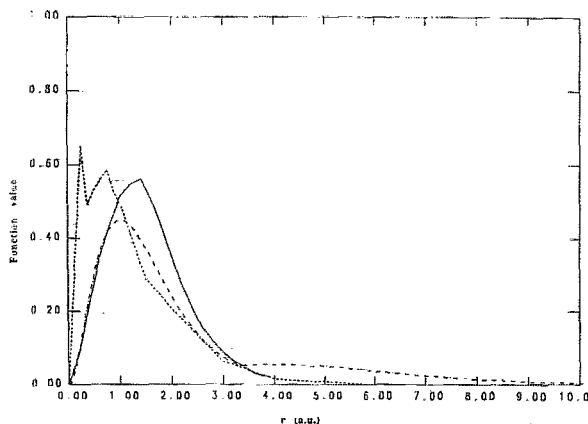


FIG. 3. Plot of the optimized guiding function (---), the original guiding function (---), and the radial density of the wave function, Eq. (11) (solid line), for the H_2 ground state. The area of each curve has been normalized to one.

TABLE IV. The energy of the H_3^+ ground state (in hartrees). N is the number of configurations. The system is in an equilateral triangle configuration with $R = 1.65$ a.u.

N	Ref. 13 ^a	Optimized $g(r)$
16 000	-1.343 586 5 ± 0.000 240 0	-1.343 847 7 ± 0.000 178 6
64 000	-1.343 660 7 ± 0.000 122 7	-1.343 720 4 ± 0.000 089 1
256 000	-1.343 762 7 ± 0.000 061 8	-1.343 739 1 ± 0.000 045 1
Frye <i>et al.</i> (Ref. 29)		-1.343 827 9 ^b
Alexander <i>et al.</i> (Ref. 27)		-1.343 822 0 ^c
Anderson (Ref. 30)		-1.343 835(1) ^d

^aVariance functional minimized using $g(r) = 12.371r^2e^{-2r}$ when $r = [0, 1]$ and $g(r) = \text{maximum}(12.371r^2e^{-2r}, 91.410r^2e^{-4r})$ when $r = [1, 100]$.

^bLinear Gaussian-type geminals calculation.

^c700 random-tempered Gaussian-type geminals.

^dDiffusion Monte Carlo.

$$\Psi_t = (1 + P_{12})(1 + P_{ABC}) \sum_{k=1}^{120} a_k [e^{-\alpha r_{1A} - \beta r_{2B}} r_{12}^{b_k} \times (r_{1A} - r_{1B})^{c_k} (r_{1A} + r_{1B})^{d_k} (r_{2A} - r_{2B})^{e_k} (r_{2A} + r_{2B})^{f_k} \times (r_{1C} - r_{2C})^{g_k} (r_{1C} + r_{2C})^{h_k}], \quad (15)$$

where A , B , and C are the positions of the three nuclei and P_{ABC} is the operator which permutes their positions. Table IV shows that optimizing the guiding function for this system decreases the statistical error by about 33%. Although the expectation value is slightly higher than the previous value, it is the same within the error. During the optimization process the value of N_{eff} rose from 54% for the previous guiding function to 73% for the optimized function. A plot of these functions is given in Fig. 4 together with $P(r)$. Like the H_2 ground state, the optimized guiding function for this system has a peak around 0.7 a.u. which places a large number of configurations in the center of the molecule and a smaller number at large r . We also see that the radial density of the wave function has its maximum further out than the optimized function does. This proba-

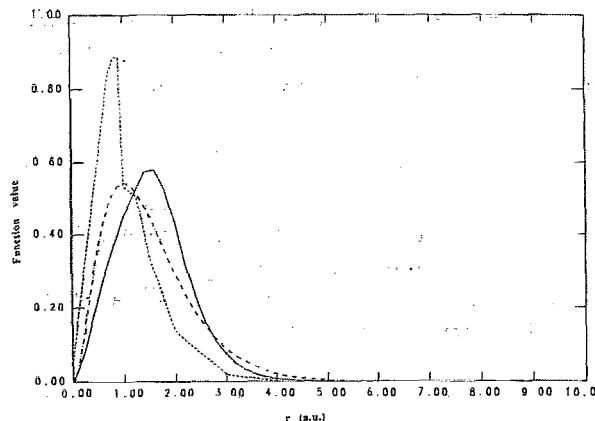


FIG. 4. Plot of the optimized guiding function (---), the original guiding function (---), and the radial density of the wave function, Eq. (11) (solid line), for the H_3^+ ground state. The area of each curve has been normalized to one.

bly reflects the contributions to the density from the other atoms in the molecule.

IV. EXCITED-STATE RESULTS

For excited states, the procedure outlined in Sec. II remains basically the same except that we now choose to explicitly orthogonalize these states with respect to the ground state. To do this, the optimization functional of the wave function must be changed from Eq. (6) to

$$\frac{\sum_i [H\Psi_i(\mathbf{x}_i) - E_{in}\Psi_i(\mathbf{x}_i)]^2 \frac{\Psi_i(\mathbf{x}_i)^2}{w(\mathbf{x}_i)^2}}{\left[\sum_i \frac{\Psi_i(\mathbf{x}_i)^2}{w(\mathbf{x}_i)} \right]^2} + \lambda \frac{\left[\sum_i \frac{\Phi_i(\mathbf{x}_i)\Psi_i(\mathbf{x}_i)}{w(\mathbf{x}_i)} \right]^2}{\sum_i \frac{\Phi_i(\mathbf{x}_i)^2}{w(\mathbf{x}_i)} \sum_i \frac{\Psi_i(\mathbf{x}_i)^2}{w(\mathbf{x}_i)}}. \quad (16)$$

Here Φ_i is the optimized wave function for the ground state and Ψ_i is the trial wave function for the first excited state. The second term explicitly orthogonalizes the first excited state to the ground state. Clearly, this expression can be generalized to even higher excited states. In Table V we list our results for the first excited singlet (2^1S) state of helium. The form of the trial wave function was taken to be the same as for the singlet ground state. During the optimization process we used the final optimized wave function and guiding function, a set of 4000 configurations, and set $E_{in} = -2.146$ hartrees. A mixing parameter $\lambda = 0.001$ was found to be adequate.

The new guiding function found with this method gives a slightly better error than the original one did. Although the expectation value is slightly lower than the previous value, it is the same within the error. During the optimization process the value of N_{eff} dropped from 42% for the previous guiding function to 27% for the optimized function. A plot of these functions is given in Fig. 5 along

TABLE V. The energy of the helium 2^1S state (in hartrees). N is the number of configurations.

N	Ref. 13 ^a	Optimized $g(r)$
16 000	$-2.145\,821\,7 \pm 0.000\,072\,5$	$-2.145\,925\,4 \pm 0.000\,040\,8$
64 000	$-2.145\,871\,2 \pm 0.000\,033\,8$	$-2.145\,894\,3 \pm 0.000\,020\,4$
256 000	$-2.145\,884\,8 \pm 0.000\,017\,1$	$-2.145\,894\,3 \pm 0.000\,010\,5$
1 024 000	$-2.145\,880\,6 \pm 0.000\,008\,6$	$-2.145\,887\,0 \pm 0.000\,005\,2$
Pekeris (Ref. 31)		$-2.145\,974\,037^b$
Thakkar and Smith (Ref. 22)		$-2.145\,973\,824^c$
Drake (Ref. 23)		$-2.145\,974\,046\,054\,143^d$
Reynolds <i>et al.</i> (Ref. 32)		$-2.144\,93(7)^e$

^aVariance functional minimized using $g(r) = 23.404r^2e^{-3r}$ when $r = [0,1]$ and $g(r) = \text{maximum}(23.404r^2e^{-3r}, 0.042r^4e^{-r})$ when $r = [1,100]$. Orthogonality with the ground state was not imposed.

^b615 functions in perimetric coordinates.

^c55 random-tempered Slater-type geminals.

^d616 Hylleraas functions with multiple sets of exponents.

^eDiffusion Monte Carlo.

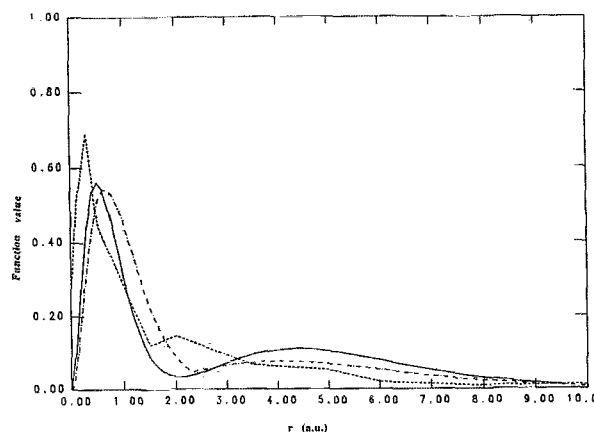


FIG. 5. Plot of the optimized guiding function (---), the original guiding function (---), and the radial density of the wave function, Eq. (11) (solid line), for the He 2^1S state. The area of each curve has been normalized to one.

with the probability of finding an electron at a distance r from nucleus A . Like the singlet ground state, the optimized guiding function differs from the original one mostly at small r and very large r .

The next system we consider is the first singlet excited ($B^1\Sigma_u^+$) state of H_2 . As in Ref. 13 we optimized the non-linear parameters with 16 000 configurations in order to get a sufficiently high N_{eff} and E_{in} was set to -0.7058 hartrees. This state is a much more interesting test of Eqs. (8) and (16) than the helium 2^1S state since our earlier variance functional optimization produced an expectation value which was significantly below that produced by Kolos and Wolniewicz.¹⁶ This apparent breakdown in the variational Monte Carlo method occurred because our final wave function was not a pure eigenfunction but rather a mixture of both ground and excited states. To some extent this problem is present in all Monte Carlo calculations which do not explicitly orthogonalize ground and excited states in the minimization step. In Ref. 13 we removed this mixing by performing a Rayleigh–Ritz type calculation *a posteriori* (because our helium 2^1S wave function was relatively pure, a Rayleigh–Ritz calculation with it produced no noticeable change). When Eq. (16) is used to optimize the trial wave function, however, we get a much better representation of this state. In Table VI we show that the

TABLE VI. The energy of the H_2 singlet excited ($B^1\Sigma_u^+$) state (in hartrees). N is the number of configurations. The internuclear distance is 1.4 a.u.

N	Ref. 13 ^a	Optimized $g(r)$
64 000	$-0.672\,963\,5 \pm 0.001\,578\,4$	$-0.691\,407\,8 \pm 0.000\,220\,6$
256 000	$-0.673\,581\,5 \pm 0.000\,791\,3$	$-0.691\,479\,9 \pm 0.000\,110\,8$
Kolos and Wolniewicz (Ref. 16)		$-0.705\,772\,684^b$

^aVariance functional minimized using $g(r) = 12.371r^2e^{-2r}$ when $r = [0,1]$ and $g(r) = \text{maximum}(12.371r^2e^{-2r}, 0.042r^4e^{-r})$ when $r = [1,100]$. Orthogonality with the ground state was imposed *a posteriori* using a Rayleigh–Ritz method.

^b249 explicitly correlated functions in elliptical coordinates.

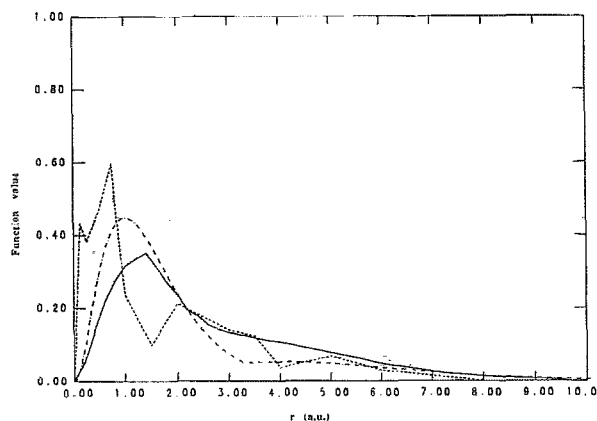


FIG. 6. Plot of the optimized guiding function (---), the original guiding function (---), and the radial density of the wave function, Eq. (11) (solid line), for the H_2 singlet excited ($B^1\Sigma_u^+$) state. The area of each curve has been normalized to one.

combination of explicitly orthogonalizing the trial wave function with respect to the ground state and optimizing the guiding function decreases the variance by over a factor of 7 compared to our earlier Rayleigh–Ritz result and produces a much better energy. During the optimization process the value of N_{eff} rose from 10% for the previous guiding function to 23% for the optimized function. A plot of these functions is given in Fig. 6 together with $P(r)$. The optimized guiding function puts many more configurations at small and some medium values of r .

V. PROPERTIES

In Ref. 14 we calculated a number of properties for several simple systems. These properties were in good agreement with the literature but their statistical errors were much larger than for the total energy. In part, this is because of cancellations between the kinetic energy and the potential energy which leave $H\Psi_i/\Psi_i$ relatively constant (this is why minimizing the variance does so well) and in part because we have chosen a guiding function which is optimal for such calculations. In principle, calculating the expectation value of some operator A is no different from calculating the expectation value of the Hamiltonian

$$\langle A \rangle = \frac{\sum_i (\Psi_i A \Psi_i / w_i)}{\sum_i (\Psi_i^2 / w_i)}, \quad (17)$$

$$\sigma_A^2 = \frac{\sum_i [(A\Psi_i - \langle A \rangle \Psi_i)^2 \Psi_i^2 / w_i^2]}{\left[\sum_i (\Psi_i^2 / w_i) \right]^2}. \quad (18)$$

We now show that guiding functions can also be optimized so as to minimize the error in these expectation values. Given a previously computed trial wave function, Ψ , the guiding function for such a property is determined by minimizing the functional

$$\frac{\sum_i \frac{[A\Psi_i(\mathbf{x}_i) - \langle A \rangle \Psi_i(\mathbf{x}_i)]^2 \Psi_i(\mathbf{x}_i)^2}{w(\mathbf{x}_i)}}{\left[\sum_i \frac{\Psi_i(\mathbf{x}_i)^2}{w(\mathbf{x}_i)} \right]^2}. \quad (19)$$

TABLE VII. Selected properties (in a.u.) of the H_2 ground state evaluated using 256 000 configurations. $R=1.4$ a.u. along the z axis.

Property	Ref. 14 ^a	Energy optimized $g(r)$	Property optimized $g(r)$	Literature
E	-1.174 44(3)	-1.174 45(2)	-1.174 45(2)	-1.174 475 668 ^b
Q_4 ^c	0.27(4)	0.31(5)	0.31(4)	0.2826 ^d
p^4	6.5(2)	6.59(9)	6.63(2)	6.6245 ^e
$\langle \delta(r_{1A}) \rangle$	0.459(1)	0.460(1)	0.4600(9)	0.46015 ^e

^aVariance functional minimized using $g(r)=12.371r^2e^{-2r}$ when $r=[0,1]$ and $g(r)=\text{maximum}(12.371r^2e^{-2r}, 0.042r^4e^{-r})$ when $r=[1,100]$.

^bReference 25.

^cThe electric hexadecapole moment for this system is defined as $Q_4=R^4/8 - \langle x_1^4 + x_2^4 + z_1^4 + z_2^4 - 6(x_1^2z_1^2 + x_2^2z_2^2) \rangle$.

^dReference 33.

^eReference 34.

Such calculations should be performed (if possible) over the same set of configurations as was used to optimize the trial wave function. This helps to correlate the errors. Once the guiding function for a property is determined then its expectation value is calculated as before, i.e., Eq. (17) is evaluated.

In Table VII we show how this method effects the accuracies of several properties of H_2 . We see immediately that a guiding function optimized to minimize the variance in the total energy noticeably improves the quality of most of the other expectation values. For the properties Q_4 and $\langle \delta(r_{1A}) \rangle$ we see that optimizing individual guiding functions produces only a slight improvement in the variance of these quantities. For the expectation value of p^4 , however, it drops the variance by a respectable amount. This is perhaps to be expected since the functional form of this quantity is very sharply peaked near the origin. In fact, our optimized guiding function reflects this behavior and places a large number of configurations in this region.

VI. CONCLUSIONS

We have shown that optimizing the guiding function in a biased-selection Monte Carlo calculation can decrease the variance of an expectation value. Although a lower variance can also be obtained by simply using a larger number of configurations, we have noticed that the expectation value itself converges more smoothly when an optimized guiding function is used. In general, we have found that a guiding function which minimizes the statistical error in the total energy is also good for most other properties (e.g., r_{12}^2) because they depend heavily on those regions where the error in the trial wave function is large. Some properties (e.g., p^4) have such a radically different functional form that one can obtain a large decrease in the variance by optimizing a separate guiding function.

Although our guiding functions are by definition one dimensional and our wave functions are (at most) six dimensional, our calculations of the radial density show that a good guiding function should be more diffuse than the trial wave function and more strongly peaked near each atomic nucleus. This suggests that those studies which use the same function for both the guiding function and the

trial wave function do not obtain the optimum expectation value or variance from a given functional form. Our results provide no information about which functional form will most closely approximate the exact wave function.

To compute excited states, we have proposed a method of keeping excited states orthogonal to all lower states during the optimization process compared to performing a Rayleigh–Ritz calculation *a posteriori*. The former requires more CPU time during the minimization step because both the excited-state wave function and all lower states must be repeatedly evaluated whereas the latter evaluates only a single trial wave function. During the expansion step, however, the Rayleigh–Ritz method must simultaneously evaluate all trial wave functions while our new method needs only the current one (since orthogonality has already been imposed). In general, this new method is easier to use and will remove the bulk of all lower states from an excited state, thereby preventing the mixing of lower states with higher states so as to approximate the excited-state energy. In most situations this method will produce a better estimate of the excited-state wave function than the Rayleigh–Ritz method. We believe, however, that the only way to be certain that an excited state is rigorously orthogonal to all lower-state approximate wave functions is to perform a Rayleigh–Ritz calculation.

Additional improvements to the calculation of expectation values can be obtained by various changes to the mechanism we use to generate configurations.^{4,6} For now, however, we prefer to keep this method as simple as possible. We have used the current algorithm to examine four electron systems and find that it performs adequately ($N_{\text{eff}} \sim 20\%$). We believe that further changes will not be needed to study systems with up to ten electrons (where $N_{\text{eff}} \sim 1\%$) until the inaccuracies in our calculations caused by the form of the trial wave function are corrected.

ACKNOWLEDGMENTS

We would like to thank the staff of the Northeast Regional Data Center for their support in running our program on the University of Florida IBM 3090. This work has been supported in part by the Defense Advanced Project Research Agency through Grant No. N00014-89-J-3030 administered by the Office of Naval Research. We also gratefully acknowledge support by the National Science Foundation through Grant No. PHY-8608155 to J. D. M. and by a grant for the Institute for Theoretical Atomic and Molecular Physics at the Harvard-Smithsonian Center for Astrophysics.

APPENDIX: THE GUIDING-FUNCTION OPTIMIZATION FUNCTIONAL

The variance depends on the guiding function through the factor $1/w(\mathbf{x}_i)^2$. If this term is replaced by $1/[w(\mathbf{x}_i)\omega(\mathbf{x}_i)]$, the variance functional in the limit of an infinite number of configurations can be written as

$$\frac{\sum_{i=1}^N [H\Psi_t(\mathbf{x}_i) - E_{in}\Psi_t(\mathbf{x}_i)]^2 \frac{\Psi_t(\mathbf{x}_i)^2}{w(\mathbf{x}_i)\omega(\mathbf{x}_i)}}{\left[\sum_{i=1}^N \frac{\Psi_t(\mathbf{x}_i)^2}{w(\mathbf{x}_i)} \right]^2} \rightarrow \frac{\int d\tau (H\Psi_t - E_{in}\Psi_t)^2 \frac{\Psi_t^2}{\omega}}{\left(\int d\tau \Psi_t^2 \right)^2}, \quad (\text{A1})$$

where a dependence on $\omega(\mathbf{x}_i)$ remains after taking the limit. When a finite number of configurations is generated with a guiding function [and corresponding weight $w(\mathbf{x}_i)$], the variance which would result from using a new guiding function [and corresponding weight $\omega(\mathbf{y}_i)$] can be estimated by

$$\frac{\sum_{i=1}^N \frac{[H\Psi_t(\mathbf{x}_i) - E_{in}\Psi_t(\mathbf{x}_i)]^2 \Psi_t(\mathbf{x}_i)^2}{w(\mathbf{x}_i) \omega(\mathbf{y}_i)}}{\left[\sum_{i=1}^N \frac{\Psi_t(\mathbf{x}_i)^2}{w(\mathbf{x}_i)} \right]^2}, \quad (\text{A2})$$

which becomes exact as $N \rightarrow \infty$. Interestingly, this procedure will not work for the Conroy functional which is independent of $w(\mathbf{x}_i)$ for large N .¹³ A guiding function which minimizes the variance should also improve the Conroy functional, but its independence of the weight function (and thus of the guiding function) may be a hint of the strength and the weakness of this functional.

- ¹ H. Conroy, J. Chem. Phys. **41**, 1327 (1964); **41**, 1331 (1964); **41**, 1336 (1964); **41**, 1341 (1964).
- ² R. L. Coldwell, Int. J. Quantum Chem. **S11**, 215 (1977).
- ³ R. L. Coldwell and R. E. Lowther, Int. J. Quantum Chem. **S12**, 329 (1978).
- ⁴ R. L. Coldwell, Int. J. Quantum Chem. **S13**, 705 (1979).
- ⁵ R. E. Lowther and R. L. Coldwell, Phys. Rev. A **22**, 14 (1980).
- ⁶ K. A. Wilson and R. L. Coldwell, J. Comput. Phys. **61**, 483 (1985).
- ⁷ M. C. Huang, R. L. Coldwell, and M. Katoot, Nucl. Phys. B **309**, 733 (1988).
- ⁸ C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988).
- ⁹ S. A. Alexander and R. L. Coldwell, Int. J. Quantum Chem. **S23**, 169 (1989).
- ¹⁰ C. J. Umrigar, Int. J. Quantum Chem. **S23**, 217 (1989).
- ¹¹ Z. Sun, S. Y. Huang, R. N. Barnett, and W. A. Lester, Jr., J. Chem. Phys. **93**, 3326 (1990).
- ¹² K. E. Schmidt and J. W. Moskowitz, J. Chem. Phys. **93**, 4172 (1990).
- ¹³ S. A. Alexander, R. L. Coldwell, H. J. Monkhorst, and J. D. Morgan III, J. Chem. Phys. **95**, 6622 (1991).
- ¹⁴ S. A. Alexander, R. L. Coldwell, Gerrard Aissing, and Ajit J. Thakkar, Int. J. Quantum Chem. (to be published).
- ¹⁵ J. M. Hammersley and D. C. Handscomb, *Monte Carlo Methods* (Methuen, London, 1964).
- ¹⁶ W. Kolos and L. Wolniewicz, Can. J. Phys. **53**, 2189 (1975).
- ¹⁷ C. L. Pekeris, Phys. Rev. **115**, 1216 (1959).
- ¹⁸ S. Rybak, K. Szalewicz, and B. Jeziorski, J. Chem. Phys. **91**, 4779 (1989).
- ¹⁹ A. J. Thakkar and V. H. Smith Jr., Phys. Rev. A **15**, 1 (1977).
- ²⁰ J. D. Baker, D. E. Freund, R. N. Hill, and J. D. Morgan III, Phys. Rev. A **41**, 1247 (1990).
- ²¹ R. J. Harrison and N. C. Handy, Chem. Phys. Lett. **113**, 257 (1985).
- ²² A. J. Thakkar and V. H. Smith, Jr., Phys. Rev. A **15**, 16 (1977).

- ²³G. F. W. Drake, Nucl. Inst. Methods Phys. Res. Sect. B **31**, 7 (1988).
²⁴S. Zhang and M. H. Kalos, Phys. Rev. Lett. **67**, 3074 (1991).
²⁵W. Kolos, K. Szalewicz, and H. J. Monkhorst, J. Chem. Phys. **84**, 3278 (1986).
²⁶D. Frye, G. C. Lie, and E. Clementi, J. Chem. Phys. **91**, 2369 (1989).
²⁷S. A. Alexander, H. J. Monkhorst, R. Roeland, and K. Szalewicz, J. Chem. Phys. **93**, 4230 (1990).
²⁸D. L. Drake and J. B. Anderson (unpublished).
²⁹D. Frye, A. Preiskorn, G. C. Lie, and E. Clementi, J. Chem. Phys. **92**, 4948 (1990).
³⁰J. B. Anderson, J. Chem. Phys. **96**, 3702 (1992).
³¹C. L. Pekeris, Phys. Rev. **127**, 509 (1962).
³²P. J. Reynolds, R. N. Barnett, B. L. Hammond, R. M. Grimes, and W. A. Lester, Jr., Int. J. Quantum Chem. **29**, 589 (1986).
³³G. Karl, J. D. Poll, and L. Wolniewicz, Can. J. Phys. **53**, 1781 (1975).
³⁴W. Kolos and L. Wolniewicz, J. Chem. Phys. **41**, 3663 (1964).