

## Optimized Trial Wave Functions for Quantum Monte Carlo Calculations

C. J. Umrigar, K. G. Wilson, and J. W. Wilkins

*Theory Center and Laboratory of Atomic and Solid State Physics,  
Cornell University, Ithaca, New York 14853*

(Received 24 December 1987)

We present a procedure for obtaining optimized trial wave functions for use in quantum Monte Carlo calculations that have both smaller statistical errors and improved expectation values, compared to commonly used functions. Results are presented for several two-electron atoms and ions (including some excited states) and for the Be atom.

PACS numbers: 31.20.Di, 71.10.+x

Monte Carlo (MC) calculations have been shown to provide energies of small atoms and molecules that are comparable to the best configuration-interaction (CI) calculations.<sup>1,2</sup> The most commonly used trial or guiding wave functions that are used in the calculation consist of a Hartree-Fock determinant multiplied by a Jastrow correlation factor (HF-J) or a multiconfiguration self-consistent-field function (consisting of a small linear combination of determinants) multiplied by a Jastrow function (MCSCF-J). Using these wave functions, variational MC calculations typically recover about 15%–80% of the correlation energy and diffusion MC and Green's-function MC calculations recover about [80–100(±2)]% of the correlation energy.<sup>3</sup> In this Letter we present a procedure for the determination of wave functions that even in a variational MC calculation recover more than 99.99% of the correlation energy for the ground and excited states of two-electron ions, and 99% of the correlation energy for the Be atom.

Past attempts at finding improved trial wave functions have mostly consisted of adjustment of the parameters of the wave function to minimize the expectation value of the energy  $\bar{E}$ . With this method it has been possible to optimize at most a few parameters since a single calculation of  $\bar{E}$  takes a significant amount of computer time, and this must be repeated several times before an optimal set of parameters is found. Instead, we have developed a procedure wherein we minimize the variance of the *local* energy.<sup>4</sup> More specifically we find the parameters in  $\psi$  that minimize

$$\sigma_{\text{opt}}^2 = \frac{\sum_{i=1}^{N_{\text{opt}}} [H\psi(i)/\psi(i) - E_g]^2 w(i)}{\sum_{i=1}^{N_{\text{opt}}} w(i)}, \quad (1)$$

where  $w(i) = |\psi(i)/\psi_0(i)|^2$ ,  $E_g$  is a guess for the energy of the state we are interested in, and the sum is over a fixed set of configurations of the electrons samples from  $|\psi_0|^2$ .  $\psi_0$  is taken to be the best wave function available before we start the optimization procedure, usually HF-J.

The chief advantage of this procedure is that 500 to 2000 configurations are found to be sufficient—a re-

markably small number considering that some of the functional forms tried had as many as 100 free parameters. There are two reasons for this. First, the configurations over which the optimization is performed are fixed, and so we are using correlated sampling to arrive at an optimal set of parameters. Hence the difference in the  $\sigma_{\text{opt}}$ 's for two sets of values of the parameters being optimized, is much more accurately determined than the values of the  $\sigma_{\text{opt}}$ 's themselves. Second, we are performing a fit, not an integral. So, if the true wave function were representable by an  $n$ -parameter trial wave function, then only  $n$  configurations would be necessary to determine the  $n$  parameters exactly. As with all nonlinear optimization problems, it is possible to get stuck in local minima. Hence the wave functions that we have determined are not necessarily the best possible for that form of the function. However, we demonstrate that it is relatively easy to find wave functions that are much better than those commonly used.

Another advantage of our procedure is that the quantity being minimized has a known lower bound (namely zero). Also, since any eigenstate has zero variance of the local energy, it is possible to calculate excited states as well as the ground states.

The initial value of  $E_g$  in Eq. (1) is taken to be the expectation value of the energy for the initial trial wave function (HF-J or MCSCF-J). Once an improved trial wave function is found, another MC run is performed with this improved function and the expectation value of the energy from this MC run is used as the new  $E_g$ . Typically only two iterations are required to obtain a self-consistent value of  $E_g$ .

The most commonly used trial wave functions have the form

$$\psi = \sum_{n=1}^{N_{\text{det}}} (d_n D_n^\uparrow D_n^\downarrow) \prod_{i < j} \exp[a_{ij} r_{ij} / (1 + b r_{ij})],$$

where  $D^\uparrow$  ( $D^\downarrow$ ) are the up- (down-) spin Slater determinants and  $r_{ij}$  is the interelectron distance of the  $i$ th and  $j$ th electrons. This form consists of a MCSCF-like

part (consisting of a sum of determinants) and a Jastrow part (involving a product over all pairs of electrons). The orbitals in the determinants are themselves linear combinations of Slater functions times spherical harmonics:

$$\phi_i(\mathbf{r}) = \sum_{j=1}^{N_{\text{basis}}} C_{ij} r^{n_j-1} e^{-\zeta_j r} Y_{l_j m_j}(\hat{\mathbf{r}}). \quad (2)$$

The electron-electron cusp condition<sup>5</sup> implies that the Jastrow  $a_{ij}$  equal  $\frac{1}{2}$  for antiparallel-spin electrons and  $\frac{1}{4}$  for parallel-spin electrons. Having a different Jastrow  $b$  for antiparallel- and parallel-spin electrons does not result in a significantly better wave function; hence a single  $b$  is used.

As a first step in finding better wave functions we used the same form for the wave functions but optimized the  $d_n$ ,  $C_{ij}$ ,  $\zeta_j$ , and the Jastrow  $b$ . This leads to significant but not large improvements in the wave function. Next we altered the form of the wave functions. In this Letter we present results for one of the more successful forms of several that were tried. For a two-electron ion there are only three independent variables, which we can take to be  $r_{12} = |\mathbf{r}_1 - \mathbf{r}_2|$ ,  $s_{12} = r_1 + r_2$ , and  $t_{12} = r_1 - r_2$ . Hence we generalize the Jastrow correlation function to include correlations between two electrons and the nucleus,

$$\prod_{i < j} \exp \left[ \frac{P(\{a\}, r_{ij}, s_{ij}, t_{ij})}{1 + P(\{b\}, r_{ij}, s_{ij}, t_{ij})} \right],$$

where  $P$  is a complete fourth-order polynomial in  $r$ ,  $s$ ,  $t$  with coefficients  $\{a\}$  or  $\{b\}$ . Antisymmetry of the wave function under the interchange of two electrons implies that the coefficients of all odd powers of  $t$  vanish. For large  $r$ ,  $s$ ,  $t$  the correlation function tends to a constant. In the limit that the order of  $P$  is infinite, it fully describes the correlation between two electrons and a nucleus in a two-electron ion, but lacks some of the higher-order correlations in a multielectron ion. Hence we ex-

pect to see, and do see, much larger improvements for two-electron ions than for multielectron ions.

The local energy  $H\psi/\psi$  in Eq. (1) diverges when an electron and a nucleus overlap ( $r=s=t$ ) or when two electrons overlap ( $r=t=0$ ) unless the wave function obeys the cusp conditions<sup>5</sup> which ensure that the infinite potential energy is canceled by the infinite kinetic energy. The reason that CI wave functions converge slowly is that they attempt to reproduce the electron-electron cusp with functions that do not have a cusp.<sup>6</sup> Our functional form is capable of satisfying the cusp conditions and therefore is more compact.

The condition for the electron-nucleus cusp is  $(\partial\psi/\partial s - \partial\psi/\partial t)|_{r=s=t} = -Z\psi$ , and that for the electron-electron cusp is  $[\partial\psi/\partial r]_{r=t=0} = \psi/2$  or  $\psi/4$  for antiparallel or parallel spins, respectively. Imposition of these conditions on our  $\psi$  results in  $17 + n_{\text{orb}}$  coupled quadratic equations where  $n_{\text{orb}}$  is the number of orbitals in the determinantal part of  $\psi$ . For maximum variational flexibility we imposed the cusp conditions approximately by including in Eq. (1) additional terms proportional to the square of the errors in satisfying the cusp conditions. By our scaling up or down the prefactor of these terms, the cusp conditions can be imposed more or less exactly.

Once an optimized wave function is obtained, a long MC calculation is carried out to calculate the expectation value of the energy  $\bar{E}$  and the fluctuation in the local energy,

$$\sigma = \left\{ \frac{1}{N_{\text{MC}}} \sum_{i=1}^{N_{\text{MC}}} [H\psi(i)/\psi(i) - E]^2 \right\}^{1/2}. \quad (3)$$

The sums are over  $N_{\text{MC}}$  configurations of electrons sampled from  $|\psi|^2$ . The value of  $N_{\text{MC}}$  was in every case between  $10^6$  and  $10^7$ . If the  $N_{\text{opt}}$  in Eq. (1) is not sufficiently large or the cusp conditions are ignored, then  $\sigma$  from Eq. (3) will be large even though  $\sigma_{\text{opt}}$  from Eq. (1) is small.

TABLE I. Performance of various wave functions for ground and excited states of two-electron ions. For each wave function, the first line gives the error in the expectation value of the energy and the second line gives the standard deviation of the local energy, in hartrees. The expected statistical error in the last digit is in parentheses.

Wave function	$1^1S H^-$	$1^1S He$	$2^3S He$	$3^3S He$	$1^1S Be^{2+}$
"Exact" energy (nonrelativistic, $\infty$ mass)	0.527 751 <sup>a</sup>	2.903 724 <sup>a</sup>	2.175 229 <sup>b</sup>	2.068 689 <sup>b</sup>	13.655 566 <sup>a</sup>
(1) HF <sup>c</sup>	0.039 821 0.31	0.042 044 0.65	0.000 978 0.054	0.000 21 0.024	0.044 267 1.4
(2) HF + Optim. (Jastrow) <sup>d</sup>	0.012 3(3) 0.082	0.015 9(6) 0.23	0.000 5(1) 0.028	0.000 10(3) 0.010	0.016(1) 0.48
(3) Optim. (Det. + Jastrow) <sup>e</sup>	0.007 9(2) 0.070	0.004 1(3) 0.14	0.000 05(5) 0.020	-0.000 04(3) 0.009(1)	0.003 5(2) 0.30
(4) Optim. (Det. + Expon. Padé) <sup>f</sup>	0.000 005(3) 0.0021	-0.000 002(4) 0.0011	-0.000 003(2) 0.00071	0.000 001(1) 0.0013	0.000 001(6) 0.0034

<sup>a</sup>Reference 7.

<sup>b</sup>Reference 8.

<sup>c</sup>0 free parameters.

<sup>d</sup>1 free parameter.

<sup>e</sup>3,4,7 free parameters for  $1^1S$ ,  $2^3S$ ,  $3^3S$  states, respectively.

<sup>f</sup>41,44,47 free parameters for  $1^1S$ ,  $2^3S$ ,  $3^3S$  states, respectively.

In Table I we present our results for a variety of two-electron ions and for several different wave functions that get progressively better. For each of the wave functions, the first line is the error in the expectation value of the energy  $\Delta E = \bar{E} - E_{\text{true}}$  and the second line is  $\sigma$ .

Wave function 1 is the HF wave function. Wave function 2 is the HF wave function with an optimized Jastrow function. For the  $1^1S$  states, both  $\Delta E$  and  $\sigma$  are reduced by approximately a factor of 3. For the  $3^3S$  states the reduction is smaller, about a factor of 2. This is to be expected since, for the  $3^3S$  states, the fact that the two electrons are in different orbitals and the antisymmetry of the wave function keep the electrons apart and so the Jastrow function plays a less important role. For the same reason,  $\Delta E$  and  $\sigma$  are smaller for the  $3^3S$  states than the  $1^1S$  states. Wave function 2 is of the type most commonly used in the literature. Hence we will compare our improved wave functions to it rather than to wave function 1.

Wave function 3 has the same form as 2, but the  $C_{ij}$ 's and the  $\zeta_j$ 's of Eq. (2) are also optimized. For each ion we increased the number of basis functions  $N_{\text{basis}}$  in Eq. (2) until no significant decrease in the value of  $\sigma$  was observed. This resulted in our using two  $1s$  basis functions for the  $1^1S$  states, two  $1s$  and one  $2s$  basis functions for the  $2^3S$  states, and two  $1s$ , one  $2s$ , and one  $3s$  basis functions for the  $3^3S$  state. Except for the very extended  $\text{H}^-$  and  $3^3S$  He states,  $\Delta E$  is improved by about a factor of 4 and  $\sigma$  by about a factor of 1.5.

As the wave function is improved, the reduction in the value of  $\Delta E$  is more rapid than the reduction in  $\sigma$ , in spite of the fact that the quantity that we are optimizing is  $\sigma$ . Hence, as the wave function is improved, the number of MC updates  $N_{\text{MC}}$  required to obtain a statistically significant value of  $\Delta E$  keeps increasing. Already for wave function 3, some systems have statistical uncertainties (given in parentheses) of comparable magnitude to  $\Delta E$ . Hence the values given are useful for the estimation of an upper bound to  $\Delta E$  rather than the value of  $\Delta E$ .

In wave function 4 the simple Jastrow correlation function is replaced by a fourth-order exponential Padé form and all the parameters are optimized. Now the correlation part of the wave function has sufficient variational freedom that we get as small a value of  $\sigma$  using a single  $\zeta$  function in the orbital of the  $1^1S$  states as we get by using a double  $\zeta$  function. The reduction in  $\sigma$ , compared to wave function 2, ranges between 7 and 200. The reduction is greater for the  $1^1S$  states than the  $3^3S$  states. Since the number of MC updates required to achieve a given uncertainty in the energy is proportional to  $\sigma^2$ , this represents a large reduction in the computer time needed.

Two of the states,  $\text{H}^-$  and  $3^3S$  He, are particularly interesting.  $\text{H}^-$  is often cited as a failure of the HF and the local-density-functional methods since they predict that  $\text{H}^-$  is not bound. However, the error in the HF en-

ergy is, in fact, slightly smaller than for the  $1^1S$  states of other two-electron ions. Our optimized wave functions yield a very accurate value for the energy, but the optimization procedure did not produce reliable wave functions until we added in a few additional configurations (besides those sampled from  $|\psi_0|^2$ ) that had electrons at a very large distance from the nucleus.

The  $3^3S$  He state is of interest since it is a true excited state, i.e., it is not the lowest state of that symmetry. (Excited states that are the lowest states of their symmetry are no harder to calculate than ground states.) Green's-function MC calculations do not give reliable energies for true excited states. However, since any eigenstate has by definition zero variance of the local energy, true excited states do not pose a special problem for our optimization scheme. The  $3^3S$  He state demonstrates this.

There have been several CI and MCSCF calculations of Be, the most accurate of which are those of Bunge,<sup>9</sup> and several MC calculations, the most accurate of which are those of Harrison and Handy.<sup>2</sup> In Table II we present our results for the Be atom. Wave functions 1-3 are similar to those of Table I. The orbitals are expanded with two  $1s$  and one  $2s$  Slater functions. The value of  $\Delta E$  of wave function 2 is in close agreement with that of Ref. 2, as it should be. Wave function 3, with optimized determinantal parameters, has a 50% smaller value of  $\Delta E$ .

The  $2p$  orbital energy of Be is nearly degenerate with the  $2s$  orbital energy and can be expected to mix in strongly. Wave function 4 is a two-configuration, four-determinant ( $1s^2 2s^2, 1s 2p_x^2, 1s 2p_y^2, 1s 2p_z^2$ ) function constructed from a two- $1s$ , one- $2s$ , one- $2p$  basis. Although  $\sigma$  improves only slightly compared with wave function 3,  $\Delta E$  improves by more than a factor of 4. The value of  $\Delta E$  is 2 times smaller than for the three-configuration (five determinant) MCSCF-J function of Ref. 2 because

TABLE II. Same as Table I, but for the Be atom.

Wave function	$1^1S$ Be
"Exact" energy (nonrelativistic, $\infty$ mass)	14.667 33(3) <sup>a</sup>
(1) HF	0.094 31
0 free parameters	1.5
(2) HF + Optim. (Jastrow)	0.060(2)
1 free parameter	0.53
(3) Optim. (Det. + Jastrow)	0.040(2)
4 free parameters	0.35
(4) Optim. (4 Det. + Jastrow)	0.009 4(9)
7 free parameters	0.34
(5) Optim. (Det. + Expon. Padé)	0.022 6(9)
44 free parameters	0.15
(6) Optim. (4 Det. + Expon. Padé)	0.000 9(3)
47 free parameters	0.092

<sup>a</sup>Extrapolated value from Ref. 9.

the parameters are not optimized in Ref. 2.

Wave function 5 is a single determinant with an exponential Padé form. Interestingly, compared to wave function 4, it has a considerably smaller  $\sigma$  but a considerably larger  $\Delta E$ . Wave function 6 consists of four determinants with an exponential Padé form. It recovers 99% of the correlation energy, more than any previous CI, MCSCF, or MC calculation of Be with the exception of the 650-configuration (43 basis functions), CI wave function of Bunge.<sup>9</sup>

We have shown that it is possible to optimize wave functions that have a large number of variational parameters. The method works for excited states as well as ground states. Although only the fluctuations are explicitly minimized, the error in the expectation value of the energy is often reduced by an even larger amount. For the simple systems treated in this paper the improvements are large. Preliminary calculations on two ten-electron systems,<sup>10</sup> Ne and water, show that significant improvements can be obtained and that  $N_{\text{opt}}$  does not increase rapidly with the number of electrons. Further improvements will hinge on the development of compact and accurate functional forms for the wave functions. The wave functions used in conventional electronic structure methods are constrained to be constructed from single-electron orbitals. MC methods, combined with the optimization scheme developed in this Letter, open up the possibility of the use of new classes of functions involving multielectron coordinates that are capable of describing the wave function more compactly and/or more exactly.

The feasibility of the optimization scheme has been demonstrated in this Letter on a few simple systems, but the scheme can be applied to any quantum MC calculation. In fact, as pointed out by Nightingale,<sup>11</sup> using the transfer matrix technique, it can also be used to improve classical MC calculations.

This work was supported by the National Science Foundation. One of us (J.W.W.) is supported by the

U.S. Department of Energy. We are grateful to Kevin Schmidt for giving us his MC program. One of us (C.J.U.) is grateful to Zachary Levine, John Rehr, Karl Runge, Jim Sethna, and Steve White for many useful discussions and to the Aspen Center for Physics for its hospitality. The computations were performed on a Gould model 9050 minicomputer and on Sun-3/50 work stations.

<sup>1</sup>Peter J. Reynolds *et al.*, J. Chem. Phys. **77**, 5593 (1982).

<sup>2</sup>R. J. Harrison and N. C. Handy, Chem. Phys. Lett. **113**, 257 (1985).

<sup>3</sup>Some typical values of the percentage correlation energy are 63% for Be<sup>2+</sup> (Ref. 2) with a HF-J function, 38% for Be (Ref. 2) with HF-J function, 79% for Be (Ref. 2) with a MCSCF-J function, 17% for H<sub>2</sub>O (Ref. 1) with a HF-J function. The corresponding percentages in a diffusion MC calculation, with the same wave functions, are 98.5%, 88%, 98%, and 84%.

<sup>4</sup>A similar procedure was developed by Robert Lynn Coldwell, Int. J. Quant. Chem. **11**, 215 (1977), and applied to the Li atom. However, only 36% of the correlation energy was recovered. The idea of minimization of the variation of the local energy goes back to J. H. Bartlett, Phys. Rev. **98**, 1067 (1955).

<sup>5</sup>T. Kato, Comm. Pure Appl. Math. **10**, 151 (1957).

<sup>6</sup>C. Schwartz, Methods Comput. Phys. **2**, 241 (1963); W. Lakin, J. Chem. Phys. **43**, 2954 (1965); B. Klahn and J. D. Morgan, J. Chem. Phys. **81**, 410 (1984); R. N. Hill, J. Chem. Phys. **83**, 1173 (1985).

<sup>7</sup>David E. Freund, Barton D. Huxtable, and John D. Morgan, Phys. Rev. A **29**, 980 (1984).

<sup>8</sup>Y. Accad, C. L. Pekeris, and B. Schiff, Phys. Rev. A **4**, 516 (1971).

<sup>9</sup>Carlos F. Bunge, Phys. Rev. A **14**, 1965 (1976).

<sup>10</sup>C. J. Umrigar *et al.*, in *Computer Simulation Studies in Condensed Matter Physics: Recent Developments*, edited by D. P. Landau and H. B. Schüttler (Springer-Verlag, Berlin, 1988).

<sup>11</sup>Peter Nightingale, private communication.