

Accuracy of electronic wave functions in quantum Monte Carlo: The effect of high-order correlations

Chien-Jung Huang

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

C. J. Umrigar

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

M. P. Nightingale

Physics Department, University of Rhode Island, Kingston, Rhode Island 02881

(Received 24 February 1997; accepted 19 May 1997)

Compact and accurate wave functions can be constructed by quantum Monte Carlo methods. Typically, these wave functions consist of a sum of a small number of Slater determinants multiplied by a Jastrow factor. In this paper we study the importance of including high-order, nucleus-three-electron correlations in the Jastrow factor. An efficient algorithm based on the theory of invariants is used to compute the high-body correlations. We observe significant improvements in the variational Monte Carlo energy and in the fluctuations of the local energies but not in the fixed-node diffusion Monte Carlo energies. Improvements for the ground states of physical, fermionic atoms are found to be smaller than those for the ground states of fictitious, bosonic atoms, indicating that errors in the nodal surfaces of the fermionic wave functions are a limiting factor.

© 1997 American Institute of Physics. [S0021-9606(97)00932-X]

I. INTRODUCTION

Optimized trial wave functions that closely approximate eigenstates of Hamiltonians are essential ingredients of accurate electronic structure calculations employing quantum Monte Carlo methods. The quality of these trial wave functions is relevant both for expectation values of physical interest and for the variance of the Monte Carlo estimators, which is a measure of the efficiency of the computation.

When one uses a variational Monte Carlo method, as the trial wave function approaches an exact eigenstate of the Hamiltonian, the energy and expectation values of quantities commuting with the Hamiltonian satisfy a zero-variance principle, i.e., the expectation values approach the exact eigenstate values, while the Monte Carlo variance goes to zero. More sophisticated forms of quantum Monte Carlo, such as diffusion Monte Carlo, attempt to project out the ground state from the trial state, in which case exact expectation values are obtained for observables that commute with the Hamiltonian, even if the trial state is not an exact eigenstate. In this case, the quality of the trial states affects only the statistical errors. However, most practical algorithms suppress admixtures of excited states completely only for nodeless wave functions, such as bosonic ground states. For trial functions with nodes, usually the fixed-node approximation¹ is made, in which case errors in the nodal surface systematically bias the expectation values. The usual variational and mixed-estimators² of observables that do not commute with the Hamiltonian always yield results for which the magnitude of the bias of the expectation values and also the statistical errors depend on the quality of the trial wave function but no zero-variance principle is satisfied in this case. In sum, for the commonly used forms of quan-

tum Monte Carlo, it is essential to employ accurate trial wave functions, in particular in the case of operators that do not commute with the Hamiltonian.

The wave functions used in electronic structure calculations employing quantum Monte Carlo usually consist of a product of a Jastrow factor and one or a sum of Slater determinants. The simplest, and possibly most commonly used wave function of this sort consists of a single determinant multiplied by a simple Jastrow factor that is a product over electron pair contributions. As regards the physics contained in the determinantal and Jastrow factors, it is generally believed that multiple determinants most efficiently incorporate near-degeneracy or non-dynamic correlation, while a Jastrow factor efficiently supplies the major portion of the dynamic correlation. It has been shown³ that a Jastrow factor that correlates two electrons and a nucleus gives much better variational energies and has much smaller fluctuations of the local energy than a Jastrow factor correlating only pairs of electrons. Given this success, one may well ask whether it is advantageous to include the next most important correlations, viz. four-body correlations of three electrons and a nucleus.

There are different measures of success that one can use to answer this question. The criteria we shall use in this paper are the reduction in the variational energy and in the fluctuations in the local energy. It is reasonable to equate a reduction in these two quantities to improvement of the quality of the wave function and to a reduction of the systematic error in expectation values of operators that do not commute with the Hamiltonian. We note that, as far as diffusion Monte Carlo is concerned, a reduction of the fluctuations in

the local energy also has the important advantage that the time-step error is usually reduced also.⁴

An alternative measure of the success of modification of the trial wave function is the improvement in the fixed-node estimate of the energy. The wave functions used in this paper were obtained by simultaneous optimization of parameters appearing in the determinantal and Jastrow factors. For fixed parameters, the former determines the location of the nodal surface and the value of the fixed-node energy. The optimization feeds back changes in the Jastrow factor to the location of the nodal surface. We observe in the work reported in this paper that this has only a small effect on the fixed-node energy. Apparently, the nodal surface changes little even for modifications of the Jastrow factor that greatly improve the wave function.

The purpose of this paper is to study the effect of incorporation of many-body correlations in the Jastrow factor. In Sec. II we present the form of the trial wave functions and discuss the cusp conditions imposed on the wave functions, which reduce greatly the number of free parameters to be optimized. To facilitate the computations, we use bases of invariants, similar to those introduced by Mushinski and Nightingale⁵ in their study of bosonic van der Waals clusters. The basis invariants employed are given in Appendix B, where we also present an algorithm for using the invariants to calculate the wave function and its first two spatial derivatives. In Sec. III we present the results for the Li, Be and Ne atoms, and also for fictitious bosonic Li, Be and Ne atoms. These latter model systems are introduced to allow us to disentangle flaws of the wave function associated with many-body correlations and the nodal surface.

II. FUNCTIONAL FORM OF THE WAVE FUNCTION

The wave functions we use have the form

$$\Psi = \sum_l d_l D_l^\dagger D_l \prod_n J_n. \quad (1)$$

D_l^\dagger and D_l are the Slater determinants of single particle orbitals for the up and down electrons respectively, and J_n is a Jastrow factor correlating n -tuples of electrons and a nucleus. The simplest, and possibly most commonly used wave function of this type contains just a single product of up- and down-spin determinants and a Jastrow factor that correlates only pairs of electrons, i.e.,

$$J_2 = \prod_{\alpha,i} \exp A(r_{\alpha i}) \prod_{i<j} \exp B_{s_{zi}+s_{zj}}(r_{ij}). \quad (2)$$

Here, the index α labels the nuclei while i and j label the electrons; $s_{zk} = \pm \frac{1}{2}$ denotes the z -component of the spin of electron k , so that the index t of B_t assumes three values, $t=0, \pm 1$, which in principle allows the correlations between electron pairs to depend on the orientation of the electron spins relative to each other and to the fixed z -component of the total electron spin. We note in this context that for given total S we always choose S_z to be maximal. Dependence of the function B_t on t allows for a better variational wave function, as judged by the variational energy and the fluctua-

tions in the local energy. More specifically, it is not possible to satisfy both the cusp conditions⁶ for parallel and anti-parallel spins unless $B_0 \neq B_{\pm 1}$. Unfortunately, however, dependence of B_t on t causes *spin contamination*,⁷ i.e., the resulting wave function is no longer an eigenfunction of the square of the total spin.

Our compromise is to allow for the minimal amount of spin dependence of the Jastrow factor required to satisfy the cusp conditions, a parsimonious approach with the added benefit of reducing the number of variational parameters. In fact, tests on wave functions with additional freedom showed only a small improvement for wave functions with three-body Jastrow functions and no improvement for wave functions with four-body Jastrow functions. To obtain the results reported here, the electron–electron part of the two-body Jastrow factor J_2 was chosen to contain spin-dependent coefficients, but the higher-order factors $J_n, n>2$, were chosen to be spin-independent. We use

$$B_t(r_{ij}) = \frac{b_t R_{ij}}{1 + b'_t R_{ij}}, \quad (3)$$

where $b_0 = 1/2$ (anti-parallel spins) and $b_{\pm 1} = 1/4$ (parallel spins); further, the

$$R_{ij} \equiv R(r_{ij}) = (1 - e^{-\kappa r_{ij}})/\kappa \quad (4)$$

are inter-particle distances scaled by the function R , as given on the right hand side of Eq. (4). The function A in Eq. (2), which has the same functional form as B , but no spin dependence, and the determinantal part of the wave function are adjusted to satisfy the electron-nucleus (e-n) cusp condition.

Both the expectation value of the variational energy and the fluctuations of the local energy can be reduced significantly by going beyond this simplest form, Eq. (2), and it has been shown³ that a large improvement can be obtained by generalizing the two-body electron-nucleus Jastrow factor to a three-body electron–electron-nucleus (e²-n) Jastrow factor

$$J_3 = \prod_{\alpha, i < j} \exp C(r_{ij}, r_{\alpha i}, r_{\alpha j}), \quad (5)$$

where again the Greek index labels the nuclei and the roman indices label the electrons. In the present paper, the function C consists of a fifth-order polynomial in the scaled inter-particle distances and terms motivated by the Fock expansion.⁸ These latter terms improve the boundary conditions satisfied by the trial wave function in reducing the dependence of the local energy on the shape of the infinitesimal triangle formed by two electrons and a nucleus, in the limit that two electrons coincide with a nucleus.⁹ The detailed form of these terms is presented in Ref. 10.

The next step is to introduce a four-body, three-electron-nucleus (e³-n) Jastrow factor

$$J_4 = \prod_{\alpha} \prod_{i < j < k} \exp D(r_{ij}, r_{jk}, r_{ki}, r_{\alpha i}, r_{\alpha j}, r_{\alpha k}). \quad (6)$$

Including these four-body terms had not been done previously in electronic structure quantum Monte Carlo calculations, but up to five-body correlations were included in

TABLE I. Total energies of Li, Be, and Ne atoms obtained in variational and diffusion Monte Carlo. For each atom, the first, second, and third rows respectively contain results using wave functions with n -body correlations respectively with $n=2$, $n=3$, and $n=4$. The number of configuration state functions in the determinantal part of the wave function is denoted by CSF. E_0 is the "exact" total energy from Ref. 15. E_c is the correlation energy. E_c^{VMC} and E_c^{DMC} are the percentages of correlation energy recovered in variational and diffusion Monte Carlo. σ_{VMC} is the root mean square fluctuation of the local energy in variational Monte Carlo. The numbers in parentheses are the statistical errors in the last digit. The last column is the efficiency of the four-body correlations as measured in Eq. (7). Energies are in Hartree atomic units.

Atom	n	CSF	E_0	E_c	E^{VMC}	E_c^{VMC} (%)	E^{DMC}	E_c^{DMC} (%)	σ_{VMC}	η
Li	2	1	-7.478 06	0.045 33	-7.474 27(4)	91.6	-7.478 01(3)	99.9	0.24	
	3	1			-7.477 88(1)	99.6	-7.478 03(1)	99.9	0.037	
	4	1			-7.477 97(1)	99.8	-7.478 03(1)	99.9	0.028	60%
Be	2	2	-14.667 36	0.094 34	-14.660 88(5)	93.1	-14.666 89(4)	99.5	0.35	
	3	2			-14.666 62(1)	99.2	-14.667 23(1)	99.9	0.089	
	4	2			-14.666 81(1)	99.4	-14.667 26(1)	99.9	0.078	31%
Ne	2	1	-128.9376	0.3905	-128.713(2)	42.5	-128.919(2)	95.2	1.9	
	3	1			-128.9008(1)	90.6	-128.9242(1)	96.6	0.90	
	4	1			-128.9029(3)	91.1	-128.9243(8)	96.6	0.88	9%

Mushinski and Nightingale's study of bosonic van der Waals clusters. Of course, inclusion of higher-body correlations in the Jastrow factor is computationally expensive, and in Ref. 5 the theory of invariants was employed to reduce the computational effort. In principle, for polynomials of high order, this approach allows one to obtain a speed-up by a factor equal to the number of elements in the particle permutation symmetry group associated with the r_{ij} and $r_{\alpha i}$ in the highest-order correlations included in the Jastrow factor. In practice, the speed-up is often considerably smaller and depends on many variable details of the computation. For further details we refer to Appendix A and Ref. 11.

As mentioned above, near-degeneracy correlations such as occur in the case of Be discussed below can be accounted for efficiently^{3,10,12,13} by inclusion of additional determinants of low-lying orbitals in Eq. (1). In this work, we use four determinants (two configuration state functions) for Be and one determinant for Li and Ne.

III. RESULTS AND DISCUSSION

A. Ground state of Li, Be, and Ne

In Table I, we show the variational Monte Carlo energies and the standard deviation of the local energies of Li, Be, and Ne atoms for wave functions containing four-body correlations and compare them with the results for two-body and three-body correlated wave functions used in our earlier work. As is well-known³ by now, inclusion of the three-body correlations results in a large improvement in the energy and a large reduction in the standard deviation of the local energy. For Li, Be and Ne, 95%, 89% and 84% of the correlation energy, missing in the two-body wave functions, is recovered while the standard deviation of the local energy is reduced by factors of 6, 4 and 2 respectively. Although not shown in the table, a further advantage accrues from the fact that the auto-correlation time of the local energies sampled in variational Monte Carlo¹⁴ and in diffusion Monte Carlo are somewhat reduced. In the present work, we observe from

Table I that inclusion of the four-body correlations also results in a significant but smaller improvement in both the variational energies and the fluctuations of the local energy. The improvement gets smaller with increasing atomic number and is disappointingly small for Ne.

In Table I, we also show the diffusion Monte Carlo energies. Since the nodes of the wave functions are determined by the determinantal part of the wave functions only, it is clear that if the same determinants are used in two wave functions, the fixed-node diffusion Monte Carlo energies must be the same to within statistical error. However, we re-optimized also the determinantal part of the wave functions when we added the higher-body terms in the Jastrow exponents, since we had hoped that the additional freedom would allow the determinantal part of the wave function to attain a more optimal nodal structure. However, we find only a small improvement in the diffusion Monte Carlo energy upon going from the two-body to the three-body Jastrow, while the difference of the three-body and four-body cases is within the statistical error. The reader should not conclude from these results that, in diffusion Monte Carlo, improvements in the wave function resulting from improved Jastrow factors are worthless. First of all, improved Jastrow factors are likely to yield more reliable expectation values of operators that do not commute with the Hamiltonian. Secondly, even if one is interested only in the diffusion Monte Carlo energy, the improved Jastrow factor results in smaller statistical errors for a given number of Monte Carlo steps and usually also in smaller time-step errors, leading to a more reliable extrapolation to the zero time-step limit. We typically find that the time-step error is much smaller for the three-body Jastrow wave function than for the corresponding two-body Jastrow wave function. For example, the three-body Jastrow wave function for Be has a time-step error that is approximately 30 times smaller than that of the two-body Jastrow wave function. However, the time-step error is of comparable magnitude for the three-body and four-body Jastrow wave functions. The best possible Jastrow factor (one

TABLE II. Total energies of the bosonic ground states of Li, Be, and Ne obtained in variational and diffusion Monte Carlo. For each atom, the first and second rows respectively contain results using wave functions with n -body correlations respectively with $n=3$ and $n=4$. E^{VMC} and E^{DMC} are the variational and diffusion Monte Carlo energies. \bar{E}^{DMC} is the average of the two E^{DMC} values which should be identical except for statistical errors. σ_{VMC} is the root mean square fluctuation of the local energy in variational Monte Carlo. The numbers in parentheses are the statistical errors in the last digit. The last column is the efficiency of the four-body correlations as measured by η in Eq. (7) using \bar{E}^{DMC} . Energies are in Hartree atomic units.

Atom	n	E^{VMC}	E^{DMC}	$E^{\text{VMC}} - \bar{E}^{\text{DMC}}$	σ_{VMC}	η
Li	3	-8.673 920(1)	-8.673 934(1)	0.000 013	0.018	
	4	-8.673 930(1)	-8.673 932(1)	0.000 003	0.011	77%
Be	3	-19.274 357(2)	-19.274 387(2)	0.000 030	0.035	
	4	-19.274 375(2)	-19.274 387(4)	0.000 012	0.021	60%
Ne	3	-266.284 11(2)	-266.284 39(2)	0.00030	0.21	
	4	-266.284 26(2)	-266.284 42(4)	0.00015	0.14	50%

that includes all-body correlations and has an infinite-order polynomial) would have a variational Monte Carlo energy that equals the fixed-node diffusion Monte Carlo energy. Hence, we measure the efficiency of the four-body contributions to the Jastrow factor by

$$\eta = \frac{E_{\text{VMC}}^{3\text{-body}} - E_{\text{VMC}}^{4\text{-body}}}{E_{\text{VMC}}^{3\text{-body}} - E_{\text{DMC}}} \quad (7)$$

The percent efficiency rapidly decreases from 60% for Li to 31% for Be to 9% for Ne; similarly, the reduction in the root mean square fluctuations of the local energy decreases from 24% for Li to 12% for Be to 2% for Ne. We attribute these somewhat disappointing results to flaws of the nodal surface of the trial wave function, which the four-body interaction Jastrow factor is not designed to correct. That is, the approximate, fixed-node, diffusion Monte Carlo wave function has discontinuous derivatives almost everywhere across the nodal surface. On the other hand, a Jastrow factor expressed in terms of inter-particle coordinates, will have non-analyticities only at the $(3N-3)$ -dimensional surface where particles coincide, which constitutes only a vanishingly small fraction of the entire $(3N-1)$ -dimensional nodal surface. Hence, over most of the nodal surface, we are attempting to describe a non-analytic function as a finite sum of analytic functions and we expect the convergence to be slow.

To test the validity of the above argument, we also performed calculations for the nodeless, bosonic ground states of the Hamiltonians of the same atoms. From Table II we see that the efficiency there is considerably greater, 77% for Li, 60% for Be and 50% for Ne. Also, the improvement in the fluctuations of the local energy is considerably larger than in the fermionic case.

This supports our conjectured explanation in terms of the nodal surface. However, we do find it considerably easier to optimize the bosonic than the fermionic wave functions. As a result, we cannot completely rule out the possibility that the

fermionic wave functions have been optimized to local minima and that considerably better local minima exist, although we consider it unlikely.

In order for the bosonic wave functions to be exact, it is necessary to include all-body correlations and to go to infinite polynomial order. However, it is expected that the low body-order correlations will be the most important ones. As mentioned in Section II, in the present paper we used fifth-order polynomials for the e^2 - n and the e^3 - n correlations. For bosonic Li, since the e^3 - n correlated wave function contains all body-order correlations, one would expect to obtain the exact ground state energy using variational Monte Carlo energy, provided that one uses polynomials of sufficiently high order. From Table II we observe that, in variational Monte Carlo, the fifth-order polynomial e^3 - n correlations account for 77% of the correlation energy missing in the e^2 - n wave function. It appears to be necessary to go to higher polynomial-order to recover the remaining 23%. In the case of bosonic Ne, the e^3 - n correlations account for 50% of the missing energy in the e^2 - n wave function, in support of our expectation that the low-body order correlations are the most important ones. In fact, in view of our observations on bosonic Li, it seems likely that, had we employed higher than fifth-order polynomials, we would have found that the e^3 - n correlations really account for more than 50% of the missing energy in the e^2 - n wave function.

For any given wave function it is relevant to ask whether further improvements can be made most economically by increasing the variational freedom of the determinantal part of the wave function (either by increasing the number of single-particle basis functions or by increasing the number of determinants) or by improving the Jastrow part of the wave function (either by increasing the polynomial-order or the body-order). It is clear that if we follow the former route to the limit of a complete basis of Slater determinants then the exact result can be obtained, even without a Jastrow factor. However, the rate of convergence would be exceedingly slow, because the Slater determinants lack singularities present in the wave function, such as the cusps at electron-electron coincidence points. These cusps can already be built into the wave function at the level of the two-body correlations by including a Jastrow factor. Incorporating three-body correlations is clearly very advantageous, but the results of this paper show that inclusion of four-body correlations may not be the most economical next step to further improvement of the wave functions, at least for the heavier systems. Instead, it may be preferable to include more determinants in the wave function. Once a sufficiently large number of determinants have been included, it seems likely that it may again become more economical to improve the Jastrow part by including the four-body correlations. Such explorations are needed in order to fully exploit the flexibility that quantum Monte Carlo offers over conventional quantum chemistry methods for the construction of accurate, yet relatively compact, wave functions.

TABLE III. Number of monomials of degree p in d inter-particle distances correlating n particles.

n	d	Number of terms of polynomial order					Total
		$p=1$	$p=2$	$p=3$	$p=4$	$p=5$	
3	3	3	6	10	15	21	55
4	6	6	21	56	126	252	461
5	10	10	55	220	715	2002	3002
6	15	15	120	680	3060	11 628	15 503

ACKNOWLEDGMENTS

This work is supported by the Office of Naval Research and NSF Grant No. DMR-9214669. We thank Claudia Filippi for useful discussions. The calculations were performed on the IBM SP2 computer at the Cornell Theory Center.

APPENDIX A: INVARIANTS: COMBINATORICS

We express the exponent of the Jastrow factor as a polynomial in (scaled) inter-particle coordinates. For an N -particle system there are $d = N(N-1)/2$ inter-particle distances. The number of monomial terms of degree p in the polynomial is $\binom{d+p-1}{p}$, which grows asymptotically as p^{d-1} for large p and as $d^p = N^{2p}$ for large N . The number of monomials of degrees 1 through 5 are shown in Table III for the three- to six-body correlations. If all N particles are different then each additional monomial adds a variational coefficient but if some of the N particles are identical, terms that are equivalent because of symmetry must have the same coefficients. In Table IV we show the number of distinct coefficients required to take into account correlations between two to four electrons and a nucleus.

It is apparent that exchange symmetry results in a considerable reduction of the number of variational coefficients, but if one simply computes the symmetric polynomials as general polynomials with constrained coefficients, one does not reduce the computational effort, which still requires a number of elementary arithmetic operations on the order of the number of coefficients of the general polynomial. On the other hand, following Ref. 5, one can speed-up the computation by using results of the theory of invariants, i.e., one can rewrite symmetric polynomials as unconstrained polynomials in new variables, viz., symmetrized sums of monomials forming a finite basis of invariants.

TABLE IV. Number of symmetrized monomials of degree p in d inter-particle distances correlating n particles; symmetrization is with respect to electron interchange. Spin-up and spin-down electrons are treated as being identical.

n	d	Number of terms of polynomial order					Total
		$p=1$	$p=2$	$p=3$	$p=4$	$p=5$	
3 (e^2 -n)	3	2	4	6	9	12	33
4 (e^3 -n)	6	2	6	14	28	54	104
5 (e^4 -n)	10	2	7	20	53	125	207

TABLE V. Number of free parameters associated with terms of degree p taking into account the reduction from imposing the cusp conditions. Spin-up and spin-down electrons are treated as being identical. The asterisk for $p=1$ indicates that the number of free parameters in the Jastrow factor is one rather than zero because the e-n cusp condition is satisfied by fixing one of the parameters in each of the orbitals rather than by fixing one of the parameters in the Jastrow factor.

n	d	Number of terms of polynomial order					Total
		$p=1$	$p=2$	$p=3$	$p=4$	$p=5$	
3 (e^2 -n)	3	1*	2	4	7	10	23
4 (e^3 -n)	6	1*	2	5	13	31	51

The decrease of computational effort is a consequence of the fact that the basis invariants themselves have to be computed only once, whereupon a number of arithmetic operations equal to the number of coefficients of the symmetric polynomial is required to complete the computation.¹⁶ If one is dealing with a system with N particles one gets the full benefit of the approach by using invariants associated with the group associated with exchange of *all* identical particles. One problem of this approach is that the basis invariants become difficult to construct, but a more fundamental problem is that this approach automatically incorporates all N -body correlations, in spite of our expectation that n -body correlations become rapidly less important as n increases. We have therefore implemented a hybrid approach which consists of using n -body invariants ($n=3$ or $n=4$ for the results reported in this paper) and symmetrizing the resulting expressions over all possible $\binom{N}{n-1}$ choices of the electrons. A method of construction of these invariants and further details will be published elsewhere.¹¹

In addition to the above symmetry considerations concerning the reduction of the number of free variational parameters, we mention that imposition of the cusp conditions⁶ has the additional advantage of ensuring that the local energy is finite at particle coincidences. Table V displays the number of free variational coefficients, after imposition of the cusp conditions, when two electrons and a nucleus (e^2 -n) and three electrons and a nucleus (e^3 -n) are correlated. Comparison of Tables IV and V shows that a large reduction in the number of variational coefficients is achieved.

APPENDIX B: BASIS INVARIANTS

In this appendix we discuss the computation of the polynomials in the Jastrow factor and we list the invariants used in our computations.

The exponent of the generalized Jastrow factor is of the following form

$$P(p_1, p_2, \dots, p_I) = \sum_{j_1, j_2, \dots, j_I} c_{j_1, j_2, \dots, j_I} p_1^{j_1} p_2^{j_2} \cdots p_I^{j_I}, \quad (\text{B1})$$

which is a polynomial of the I basis polynomials p_1, p_2, \dots, p_I . This polynomial can be computed efficiently using a scheme derived from the recursive, multi-variate generalization of Horner's rule. That is, for a polynomial

$c_0 + c_1 p + c_2 p^2 + c_3 p^3 + \dots$ in one variable p , the computation is performed as suggested by the expression $c_0 + p\{c_1 + p[c_2 + p(c_3 + \dots)]\}$. For many variables, one considers P to be a polynomial in p_I and one applies the one-variable rule for its evaluation. The coefficients of this polynomial in p_I are polynomials in p_1, \dots, p_{I-1} . These in turn are computed recursively with the same scheme.

We use a variant of this algorithm designed to evaluate the monomials in P separately, rather than just P itself. The reason is that we use a fixed set of configurations for optimizing the trial wave functions, so that the monomials comprising the above polynomial do not change during the optimization provided that the scale factor κ in Eq. (4) is kept constant. We note parenthetically that fixing the value of κ does not result in a large reduction in the variational freedom for a significant range of values of κ around the optimal value. Hence, it is computationally efficient to save the values of the monomials for each configuration of the sample used for the wave function optimization, and to evaluate the value of the polynomial P for each configuration as a dot product of this constant vector of monomials with the varying coefficient vector c_{j_1, j_2, \dots, j_I} .

Our variant of Horner's rule is as follows: We start with the basis invariants p_1, \dots, p_I and regard them as the basic variables and therefore of degree one. At the second stage, we construct all possible polynomials quadratic in the basis invariants by multiplying all invariants of first degree by p_1 , then multiply all except p_1 by p_2 , all except p_1 and p_2 by p_3 and so on and so forth.

At stage n of the calculation we construct monomials of degree n in the basis invariants p_1, \dots, p_I from those of degree $n-1$ obtained at the previous stage by multiplying (1) all the monomials of degree $n-1$ by p_1 ; (2) all except those descended from p_1 by p_2 ; (3) all except those descended from p_1 and p_2 by p_3 ; etc. Hence, we obtain the following sequence of monomials:

$$v_1 = (p_1; p_2; p_3; \dots; p_I), \quad (\text{B2})$$

$$v_2 = (p_1^2, p_1 p_2, p_1 p_3, \dots; p_2^2, p_2 p_3 \dots; p_3^2 \dots; \dots; p_I^2),$$

$$v_3 = (p_1^3, p_1^2 p_2, p_1^2 p_3, \dots; p_1 p_2^2, p_1 p_2 p_3, \dots; p_1 p_3^2 \dots; \dots), \quad (\text{B3})$$

$$p_2^3, p_2^2 p_3, \dots; p_2 p_3^2 \dots; p_3^3 \dots; \dots; p_I^3), \quad (\text{B4})$$

⋮

$$v_n = (p_1^n, p_1^{n-1} p_2, \dots; p_2^n, p_2^{n-1} p_3, \dots; p_3^n, \dots; \dots; p_I^n). \quad (\text{B5})$$

Generalization of the above algorithm to the computation of the gradient and Laplacian required in the computation is straightforward. The basis invariants as functions of the scaled inter-particle distances are homogeneous polynomials of various degrees, and in practical applications one truncates the polynomial P at some chosen degree in the inter-particle distances. The above algorithm, however, will generate monomials of varying degrees in these at any step. This can be corrected by simply not constructing monomials that exceed the maximal degree.

For the case of two identical electrons (labeled by i and j) and a nucleus (labeled by α) the following set of basis invariants can be used:

$$p_1 = r_{ij}, \quad (\text{B6})$$

$$p_2 = r_{\alpha i} + r_{\alpha j}, \quad (\text{B7})$$

$$p_3 = r_{\alpha i} r_{\alpha j}. \quad (\text{B8})$$

For three electrons (denoted by subscripts i, j, k) and a nucleus (denoted by α) we used:

$$p_1 = r_{ij} + r_{ik} + r_{jk}, \quad (\text{B9})$$

$$p_2 = r_{\alpha i} + r_{\alpha j} + r_{\alpha k}, \quad (\text{B10})$$

$$p_3 = r_{ij}^2 + r_{ik}^2 + r_{jk}^2, \quad (\text{B11})$$

$$p_4 = r_{\alpha i}^2 + r_{\alpha j}^2 + r_{\alpha k}^2, \quad (\text{B12})$$

$$p_5 = r_{\alpha k} r_{ij} + r_{\alpha j} r_{ik} + r_{\alpha i} r_{jk}, \quad (\text{B13})$$

$$p_6 = r_{ij} r_{ik} r_{jk}, \quad (\text{B14})$$

$$p_7 = r_{\alpha i} r_{\alpha j} r_{\alpha k}, \quad (\text{B15})$$

$$p_8 = r_{\alpha i} r_{\alpha j} r_{ij} + r_{\alpha i} r_{\alpha k} r_{ik} + r_{\alpha j} r_{\alpha k} r_{jk}, \quad (\text{B16})$$

$$p_9 = r_{\alpha i} r_{ij} r_{ik} + r_{\alpha j} r_{ij} r_{jk} + r_{\alpha k} r_{ik} r_{jk}. \quad (\text{B17})$$

Finally, we note that the above choice of basis invariants is not unique. They were constructed to yield invariants consisting of a small number of monomials each.¹¹ It should be noted that we have no proof that the nine invariants given above indeed form a complete basis in the mathematical sense, but they are complete for the fifth-order polynomials used in this paper.

¹J. B. Anderson, *J. Chem. Phys.* **65**, 4121 (1976); J. M. Moskowitz, K. E. Schmidt, M. A. Lee, and M. H. Kalos, *ibid.* **77**, 349 (1982); P. J. Reynolds, D. M. Ceperley, B. J. Adler, and W. A. Lester, *ibid.* **77**, 5593 (1982).

²D. M. Ceperley and M. H. Kalos, in *Monte Carlo Methods in Statistical Physics*, edited by K. Binder, Topics Current Phys., Vol. 7 (Springer, Berlin, Heidelberg, 1979).

³C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, in *Computer Simulation Studies in Condensed Matter Physics: Recent Developments*, edited by D. P. Landau and H. B. Schüttler (Springer-Verlag, Berlin, 1988); *Phys. Rev. Lett.* **60**, 1719 (1988).

⁴C. J. Umrigar, M. P. Nightingale, and K. J. Runge, *J. Chem. Phys.* **99**, 2865 (1993).

⁵A. Mushinski and M. P. Nightingale, *J. Chem. Phys.* **101**, 8831 (1994).

⁶T. Kato, *Commun. Pure Appl. Math.* **10**, 151 (1957).

⁷C. Huang, C. Filippi, and C. J. Umrigar (unpublished).

⁸V. Fock, *Izv. Akad. Nauk. SSSR, Ser. Fiz.* **18**, 161 (1954); English translation: *D. Kngl. Norske Videnskab. Selsk. Forh.* **31**, 138–145 (1958).

⁹C. R. Myers, C. J. Umrigar, J. P. Sethna, and J. D. Morgan, *Phys. Rev. A* **44**, 5537 (1991).

¹⁰C. Filippi and C. J. Umrigar, *J. Chem. Phys.* **105**, 213 (1996).

¹¹M. P. Nightingale, M. Meierovich, and C. Huang (unpublished).

¹²R. J. Harrison and N. C. Handy, *Chem. Phys. Lett.* **113**, 257 (1985).

¹³H-J Flad, M. Caffarel, and A. Savin, in *Recent Advances in Quantum Monte Carlo Methods*, edited by W. A. Lester, Jr. (World Scientific, Singapore, 1997).

¹⁴C. J. Umrigar, Phys. Rev. Lett. **71**, 408 (1993).

¹⁵E. R. Davidson, S. A. Hagstrom, S. J. Chakravorty, V. M. Umar, and C. F. Fischer, Phys. Rev. A **44**, 7071 (1991); S. J. Chakravorty, S. R. Gwaltney, E. R. Davidson, F. A. Parpia, and C. F. Fischer, Phys. Rev. A **47**, 3649 (1993).

¹⁶Strictly speaking this statement is correct only if the basis invariants are algebraically independent. In practice, the algebraic dependences introduce only very few additional coefficients in the range of polynomial degrees of practical interest.