

# Spin contamination in quantum Monte Carlo wave functions

Chien-Jung Huang, Claudia Filippi,<sup>a)</sup> and C. J. Umrigar

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

(Received 8 September 1997; accepted 19 February 1998)

The wave function usually employed in quantum Monte Carlo (QMC) electronic structure calculations is the product of a Jastrow factor and a sum of products of up-spin and down-spin determinants. Typically, a different Jastrow factor is used for parallel- and antiparallel-spin electrons in order to satisfy the cusp conditions and thereby ensure that the local energy at electron-electron coincidence points is finite. However, when the Jastrow factor is not completely symmetric under interchange of the spatial coordinates of the electrons, the resulting wave function may not be an eigenstate of the spin operator  $\hat{S}^2$ . For the Li and Be atoms, we evaluate the spin contamination in a variety of wave functions with progressively higher body-order correlations in the Jastrow factor. The spin contamination is found to be small for all the wave functions ( $10^{-3}$ – $10^{-5}$ ), the smaller values being obtained for the more accurate wave functions. On the other hand, if we eliminate the spin contamination by employing a symmetric Jastrow (and sacrificing the parallel-spin cusp condition), the resulting wave functions typically have about 20%–40% larger root mean square fluctuations in the local energy. A wave function which both satisfies the cusp conditions and has definite spin can be constructed, but for systems with many electrons, its computational cost is higher than for the commonly used QMC wave functions. © 1998 American Institute of Physics. [S0021-9606(98)02720-2]

## I. INTRODUCTION

In conventional quantum chemistry methods such as the configuration interaction method, approximate wave functions are constructed to be antisymmetric and have the spin and space symmetry of the true wave function. However, since the wave functions are expressed as linear combinations of a finite number of determinants of single-particle orbitals, the wave functions do not satisfy the cusp conditions at electron-electron coincidence points and, consequently, have infinite local energy there. One of the advantages of quantum Monte Carlo (QMC) methods compared to the conventional quantum chemistry methods is that Jastrow factors can be employed to construct compact wave functions that satisfy the cusp conditions. On the other hand, these wave functions are not, in general, eigenfunctions of the spin operator  $\hat{S}^2$ .<sup>1</sup> A wave function which both satisfies the cusp conditions and is an eigenfunction of  $\hat{S}^2$  can be constructed, but its computational cost is substantially higher than for conventional QMC wave functions.

In the present paper, we show how one can determine the *spin contamination*,<sup>2,3</sup> i.e., the admixture of spurious spin components, in a given QMC wave function and derive expressions to separate the various spin components present in a spin-contaminated wave function. As examples, we evaluate the spin contamination in several QMC wave functions for the Li and Be atoms. All the wave functions we consider

are the product of a determinantal part and a Jastrow part but have different functional forms for the Jastrow factor. We employ Jastrow factors, including progressively higher body-order correlations.<sup>1,4</sup> All the parameters entering in the different wave functions are optimized using the variance minimization method.<sup>5,6</sup> All our wave functions that obey the cusp conditions have a very small spin contamination which gets even smaller when higher-order correlations are included in the wave function. On the other hand, the use of a symmetrical Jastrow, violating the parallel-spin electron cusp condition, yields a wave function with no spin contamination but at the price of a significant increase in the fluctuations of the local energy.

In Sec. II, we discuss the form of the QMC wave functions and in Sec. III the cusp conditions that the true wave function must satisfy. In Sec. IV, we derive formulas for separating the spin components of a wave function and for calculating the spin contamination. The quantum Monte Carlo method used for calculating the spin contamination and its effect on charge density is discussed in Sec. V. In Sec. VI, the spin contamination is calculated for a variety of Li and Be wave functions. In Appendix B, we further discuss cusp conditions and present two wave functions for which the usual prescription for the Jastrow factor fails to satisfy the cusp conditions. In Appendix C, we use the Li and Be atoms to illustrate the formulas derived in Sec. IV for evaluating the spin contamination. The forms of the QMC wave functions we used in the calculations are presented in Appendix D.

<sup>a)</sup>Present address: Department of Physics, University of Illinois at Urbana-Champaign, Urbana, Illinois 61801.

## II. QUANTUM MONTE CARLO WAVE FUNCTIONS

A wave function  $\Psi$  for an  $N$ -electron system ( $N=N_\uparrow+N_\downarrow$ ) in an  $S_z=(N_\uparrow-N_\downarrow)/2$  state can be decomposed in terms of its spin components as

$$\begin{aligned} \Psi(\mathbf{r}_1 s_1, \dots, \mathbf{r}_N s_N) &= \sum_{i=1}^K F_i(\mathbf{r}_1, \dots, \mathbf{r}_N) \zeta_i(s_1, \dots, s_N) \\ &= \mathcal{A}\{F_1(\mathbf{r}_1, \dots, \mathbf{r}_N) \zeta_1(s_1, \dots, s_N)\}, \end{aligned} \quad (1)$$

where

$$\begin{aligned} \zeta_1(s_1, \dots, s_N) &= \chi_\uparrow(s_1) \dots \chi_\uparrow(s_{N_\uparrow}) \\ &\quad \times \chi_\downarrow(s_{N_\uparrow+1}) \dots \chi_\downarrow(s_N). \end{aligned} \quad (2)$$

The sum in Eq. (1) extends over the  $K=\binom{N}{N_\uparrow} = N!/(N_\uparrow!N_\downarrow!)$  spin functions (or spin assignments)  $\zeta_i$  generated by permuting particle indices in  $\zeta_1$ , and  $\mathcal{A}$  is the total antisymmetrizer. It follows from the antisymmetry of  $\Psi$  under the interchange of particle indices that each function  $F_i$  is antisymmetric under the interchange of like-spin electrons and that the  $F_i$  are all the same except for a relabeling of the particle indices and a change in sign for odd permutations.

Since the spin functions  $\zeta_i(s_1, \dots, s_N)$  form an orthonormal set,

$$\sum_{s_1, \dots, s_N} \zeta_i(s_1, \dots, s_N) \zeta_j(s_1, \dots, s_N) = \delta_{i,j}, \quad (3)$$

the expectation value of an operator that is independent of spin is the same if we use the fully antisymmetric wave function  $\Psi$  or just one spatial function, say  $F_1$ .

It is sometimes convenient to express a wave function  $\Psi$  as a product of functions  $\Phi$  and  $J$  that are, respectively, antisymmetric and symmetric under particle exchange. Expanding each of  $\Phi$  and  $J$  in terms of its spin components, we obtain

$$\begin{aligned} \Psi(\mathbf{r}_1 s_1, \dots, \mathbf{r}_N s_N) &= \Phi(\mathbf{r}_1 s_1, \dots, \mathbf{r}_N s_N) J(\mathbf{r}_1 s_1, \dots, \mathbf{r}_N s_N) \\ &= \left( \sum_{i=1}^K f_i(\mathbf{r}_1, \dots, \mathbf{r}_N) \zeta_i(s_1, \dots, s_N) \right) \\ &\quad \times \left( \sum_{i=1}^K G_i(\mathbf{r}_1, \dots, \mathbf{r}_N) \zeta_i(s_1, \dots, s_N) \right) \\ &= \sum_{i=1}^K F_i(\mathbf{r}_1, \dots, \mathbf{r}_N) \zeta_i(s_1, \dots, s_N), \end{aligned} \quad (4)$$

where  $F_i=f_i G_i$  and we have used the relation  $\zeta_i \zeta_j = \delta_{i,j} \zeta_i$ .

The usual quantum chemistry wave functions have  $J=1$  and

$$\Phi = \sum_{i=1}^{N_{\text{det}}} d_i \mathbf{D}_i(\mathbf{r}_1 s_1, \dots, \mathbf{r}_N s_N), \quad (5)$$

where  $\mathbf{D}(\mathbf{r}_1 s_1, \dots, \mathbf{r}_N s_N)$  is a determinant of  $N$  spin orbitals. In that case,

$$\begin{aligned} F_1(\mathbf{r}_1, \dots, \mathbf{r}_N) &= \sum_i d_i \mathbf{D}_i^\uparrow(\mathbf{r}_1, \dots, \mathbf{r}_{N_\uparrow}) \\ &\quad \times \mathbf{D}_i^\downarrow(\mathbf{r}_{N_\uparrow+1}, \dots, \mathbf{r}_N). \end{aligned} \quad (6)$$

Quantum Monte Carlo wave functions employ the same form for  $\Phi$  (though often with reoptimized values of the wave function parameters) and a Jastrow factor  $J(\{r_{ij}\}, \{r_{i,\alpha}\}, \{s_i\})$  depending on the interelectron distances,  $\{r_{ij}\}$ , the electron–nuclei distances,  $\{r_{i,\alpha}\}$ , and the electron spins,  $\{s_i\}$ . The spin-assigned wave function is then

$$\begin{aligned} \Psi^{\text{QMC}} &= F_1(\mathbf{r}_1, \dots, \mathbf{r}_N) \\ &= J(\{r_{ij}\}, \{r_{i,\alpha}\}) \sum_i d_i \mathbf{D}_i^\uparrow(\mathbf{r}_1, \dots, \mathbf{r}_{N_\uparrow}) \\ &\quad \times \mathbf{D}_i^\downarrow(\mathbf{r}_{N_\uparrow+1}, \dots, \mathbf{r}_N). \end{aligned} \quad (7)$$

We point out that we have used the same symbol  $J$  to denote both the full space-spin Jastrow factor  $J(\{r_{ij}\}, \{r_{i,\alpha}\}, \{s_i\})$  and its first spin-assigned component  $J(\{r_{ij}\}, \{r_{i,\alpha}\}) \equiv G_1(\mathbf{r}_1, \dots, \mathbf{r}_N)$  but the meaning should be clear from the context.

The wave function  $\Psi$  must be an eigenstate of  $S^2$  and  $S_z$ . (For Hamiltonians with spherical or cylindrical symmetry it must also be an eigenstate of  $L^2$  and  $L_z$  or  $L_x$ , respectively. Since the spatial dependence of the Jastrow factor is expressed only in interparticle coordinates, it is invariant under spatial rotations and does not affect the spatial symmetry determined by the determinantal part.) We employ  $N_\uparrow$  spin-up functions and  $N_\downarrow$  spin-down functions in construction of the spin functions, and consequently  $\Psi$  is an eigenfunction of  $S_z$ . It is shown in Sec. IV that it is an eigenstate of  $S^2$  if the functions  $F_i$  satisfy certain linear relations. The quantum chemistry wave functions  $\Phi$  are constructed to be eigenstates of  $S^2$ . It is apparent from Eq. (4) that if the Jastrow factor  $J$  is symmetric under the exchange of just the spatial coordinates of any pair of electrons, i.e., if  $G_i$  is independent of  $i$ , then the  $F_i$  satisfy the same linear relations that the  $f_i$  satisfy and, therefore,  $\Psi$  is an eigenstate of  $S^2$  if  $\Phi$  is. However, as discussed in the next section, the commonly used forms of the Jastrow factor are chosen to satisfy cusp conditions (discussed in the next section) and are not invariant under the exchange of just the spatial coordinates of two opposite-spin electrons and, consequently, the Jastrow factor can introduce spin contamination into the wave function. In practice, we find that the spin contamination is very small for optimized wave functions. As discussed in Sec. IV, the spin symmetry can be enforced without violating the cusp conditions by projecting out a pure spin component from a contaminated wave function. However, we show in Appendix A that this increases the computational cost of evaluating expectation values of spin-independent operators from being  $\mathcal{O}(N^3)$  to  $\mathcal{O}(N^3 \times N)$ .

## III. CUSP CONDITIONS

The true wave function must satisfy a set of ‘‘cusp conditions’’ which prescribe the proper derivative discontinuity at the collision points and ensure that the divergence in the

local potential is canceled by an opposite divergence in the local kinetic energy. Kato<sup>7</sup> first rigorously derived these conditions as a general property of Coulombic systems. He showed that in the limit that two particles of masses  $m_i$  and  $m_j$  and charges  $q_i$  and  $q_j$  approach each other and all other interparticle distances remain larger than zero,

$$\left. \frac{\partial \hat{\Psi}}{\partial r_{ij}} \right|_{r_{ij}=0} = \mu_{ij} q_i q_j \Psi(r_{ij}=0), \quad (8)$$

where  $\mu_{ij} = m_i m_j / (m_i + m_j)$  is the reduced mass of the two-particle subsystem and  $\hat{\Psi}$  is the average of  $\Psi$  over an infinitesimally small sphere centered at  $r_{ij}=0$ . Pack and Byers-Brown<sup>8</sup> generalized Kato's result by solving the multiparticle Schrödinger equation in the neighborhood of a two-particle coincidence. Writing the wave function near such a collision point,  $\mathbf{r}_{ij}=0$ , as

$$\Psi = \sum_{l=0}^{\infty} \sum_{m=-l}^l f_{lm}(r) r^l Y_{lm}(\theta, \phi), \quad (9)$$

where  $r=r_{ij}$ ,  $\theta=\theta_{ij}$ ,  $\phi=\phi_{ij}$ , and the  $Y_{lm}(\theta, \phi)$  are the usual spherical harmonics and, expanding  $f_{lm}(r) = \sum_{k=0}^{\infty} f_{lm}^{(k)} r^k$ , they found for electron-electron collisions,

$$f_{lm}(r) = f_{lm}^{(0)} \left[ 1 + \frac{1}{2(l+1)} r + O(r^2) \right]. \quad (10)$$

The terms  $r^l Y_{lm}(\theta, \phi)$  in Eq. (9) are analytic functions that can be well described by the determinants appearing in Eq. (7), while the terms  $[1 + [1/2(l+1)]r]$  are nonanalytic functions of the individual electron coordinates, having cusps of magnitude  $[1/2(l+1)]$  (though they are of course analytic functions of the interparticle distance  $r$ ). The local energy at the collision point is finite provided that the cusp-condition [Eq. (10)] is obeyed for the lowest value  $l_0$  appearing in the sum in Eq. (9).

Cusp conditions are imposed on QMC wave functions by multiplying the spin-assigned wave function by an appropriate Jastrow factor that has the correct cusps. The usual practice is to employ a Jastrow factor with a cusp of 1/4 for parallel-spin electrons and a cusp of 1/2 for antiparallel-spin electrons. The former condition is always correct, while the latter is subject to some caveats, as discussed in Appendix B.

The fully antisymmetric wave function  $\Psi$  of Eq. (5) is only four times as computationally expensive as the spin-assigned wave function  $F_1$  [Eq. (7)] since the cost of evaluating the determinant of an  $N \times N$  matrix scales as  $\mathcal{O}(N^3)$ . However, once the cusp conditions are satisfied by employing a Jastrow factor, the construction of a fully antisymmet-

ric wave function is a factor of  $[N!/(N_{\uparrow}!N_{\downarrow}!)]$  more expensive: one has to explicitly perform the sum in Eq. (4) unless the Jastrow factor is independent of spin. As shown in Appendix A, the computational cost for evaluating expectation values of spin-independent operators is  $\mathcal{O}(N)$  greater for a spin-projected wave function than for the spin-assigned function  $F_1$ , and consequently it is more efficient to use spin-assigned wave functions in QMC calculations.

#### IV. SPIN CONTAMINATION

We now derive the conditions that the functions  $F_i$  [Eq. (4)] must satisfy in order for  $\Psi$  to be an eigenstate of  $\hat{S}^2$ . Since the wave functions used in QMC calculations are usually not normalized, we do not assume a normalized  $\Psi$ . Suppose  $\Psi$  is an admixture of different spin components,

$$\Psi = \sum_S \Psi_S, \quad (11)$$

where  $\Psi_S$  is the component of  $\Psi$  in the spin  $S$  subspace  $\Omega_S$ .  $\Psi_S$  can be projected out from  $\Psi$  using the spin projection operator,<sup>9</sup>

$$\Psi_S = \prod_{S' \neq S} \left[ \frac{\hat{S}^2 - S'(S'+1)}{S(S+1) - S'(S'+1)} \right] \Psi. \quad (12)$$

When there is more than one state in  $\Omega_S$ ,  $\Psi_S$  can be further decomposed into these  $N_S$  degenerate (with respect to  $\hat{S}^2$ ) states,

$$\Psi_S = \sum_{j=1}^{N_S} c_j \psi_j, \quad (13)$$

with normalized  $\psi_j$ .

Diagonalizing the operator  $\hat{S}^2$  is generally the most efficient way to perform the projection. The total spin operators  $\hat{S}_{\alpha}$  are sums of the corresponding single-electron spin components  $\hat{s}_{i,\alpha}$  for  $\alpha=x,y,z$ . The operator  $\hat{S}^2$  can be rewritten in terms of lowering and raising operators as

$$\begin{aligned} \hat{S}^2 &= \hat{S}_z^2 - \hat{S}_z + \hat{S}^+ \hat{S}^- \\ &= \hat{S}_z^2 - \hat{S}_z + \sum_{i=1}^N \hat{s}_i^+ \hat{s}_i^- + \sum_{i \neq j}^N \hat{s}_i^+ \hat{s}_j^-. \end{aligned} \quad (14)$$

The first three terms of  $\hat{S}^2$  are diagonal, while the last term is nondiagonal and exchanges the spins of a spin-up spin-down pair. Hence,

$$\Theta_{ij} \equiv \langle \zeta_i | \hat{S}^2 | \zeta_j \rangle = \begin{cases} N_{\uparrow} + S_z^2 - S_z & \text{if } i=j, \\ 1 & \text{if } \zeta_i \text{ and } \zeta_j \text{ are related by a single exchange of a spin-up spin-down pair,} \\ 0 & \text{otherwise.} \end{cases} \quad (15)$$

The  $\Theta$  matrix is symmetric and hence can be diagonalized,

$$\Theta = T^t D T, \tag{16}$$

where  $T$  is orthonormal and  $D_{ij} = \delta_{ij} S_i(S_i + 1)$ . From the orthonormality of  $T$ , we obtain

$$\begin{aligned} \Psi &= \sum_{i=1}^K F_i \zeta_i = \sum_{i=1}^K \tilde{F}_i \tilde{\zeta}_i, \\ \tilde{\zeta}_i &= \sum_{j=1}^K T_{ij} \zeta_j, \quad \tilde{F}_i = \sum_{j=1}^K T_{ij} F_j. \end{aligned} \tag{17}$$

The  $\{\tilde{\zeta}_i\}$  are the eigenstates of  $\hat{S}^2$ . Suppose  $\tilde{\zeta}_i$  has spin eigenvalue  $S(S+1)$ . If there is more than one state in  $\Omega_S$ ,  $\tilde{F}_i \tilde{\zeta}_i$  need not equal any  $c_i \psi_i$  in Eq. (13) and, in fact, need not even be antisymmetric. Nevertheless, summing over the degenerate  $\hat{S}^2$  states, we obtain the spin  $S$  component of  $\Psi$ ,

$$\Psi_S = \sum_{i=1}^{N_S} c_i \psi_i = \sum_{i=1}^{N_S} \tilde{F}_i \tilde{\zeta}_i, \tag{18}$$

which is antisymmetric. Therefore,  $\Psi_S$  can be obtained either by using the projection operator [Eq. (12)] or from Eq. (18). Since both  $\{\tilde{\zeta}_i\}$  and  $\{\psi_i\}$  are orthonormal sets, if we square Eq. (18), sum over spin variables, and integrate over the spatial coordinates, we obtain

$$\sum_{i=1}^{N_S} c_i^2 = \sum_{i=1}^{N_S} \int d^{3N} \mathbf{r} \tilde{F}_i^2. \tag{19}$$

If we start from a wave function  $\Psi$  that satisfies the cusp conditions but is not an eigenstate of  $\hat{S}^2$ , we can project out a wave function  $\Psi_S$  that both satisfies the cusp conditions and is an eigenstate of  $\hat{S}^2$ . We show in Appendix A that, in QMC calculations, the computational cost of evaluating expectation values of spin-independent operators increases from  $\mathcal{O}(N^3)$  to  $\mathcal{O}(N^3 \times N)$  if  $\Psi_S$  is used instead of  $\Psi$ .

Suppose that we are interested in constructing a trial wave function with a spin  $S$  but that  $\Psi$  contains an admixture of spin states.  $\Psi$  can be decomposed as

$$\Psi = \Psi_S + \sum_{S' \neq S} \Psi_{S'}, \tag{20}$$

where the second term is the spin contamination and, for the wave functions usually employed in quantum Monte Carlo, is much smaller than  $\Psi_S$  in magnitude.

Operating with  $\{\hat{S}^2 - S(S+1)\}$  on  $\Psi$  annihilates the spin  $S$  component of  $\Psi$ , leaving the components of  $\Psi$  with spin not equal to  $S$ ,

$$\begin{aligned} \Psi'_S &= \{\hat{S}^2 - S(S+1)\} \Psi \\ &= \sum_{S' \neq S} [S'(S'+1) - S(S+1)] \Psi_{S'}. \end{aligned} \tag{21}$$

The condition of no spin contamination can be defined as  $\Psi'_S = 0$  or, equivalently,  $\Psi_{S'} = 0$  if  $S' \neq S$ . Since  $\Psi_{S'}$  can be expressed in terms of  $\tilde{F}_i$ , the condition can also be written as

$$\tilde{F}_i = \sum_{j=1}^K T_{ij} F_j = 0, \quad \text{if } S_i \neq S, \tag{22}$$

where  $S_i$  is the spin value of  $\tilde{\zeta}_i$ . These are the linear relations that the  $F_i$ 's must satisfy for  $\Psi = \sum_{i=1}^K F_i \zeta_i$  to be a pure spin  $S$  state.

We define the spin contamination of  $\Psi$  to be the norm of the components of  $\Psi$  other than the desired spin  $S$  component divided by the norm of  $\Psi$ ,

$$\begin{aligned} \delta S^2 &\equiv \frac{\langle \Psi'_S | \Psi'_S \rangle}{\langle \Psi | \Psi \rangle} = \frac{\langle \Psi | \{\hat{S}^2 - S(S+1)\}^2 | \Psi \rangle}{\langle \Psi | \Psi \rangle} \\ &= \frac{\sum_{i, S_i \neq S} [S_i(S_i+1) - S(S+1)]^2 \int d^{3N} \mathbf{r} \tilde{F}_i^2}{\sum_{i=1}^K \int d^{3N} \mathbf{r} \tilde{F}_i^2} \\ &= \frac{\sum_{i, S_i \neq S} [S_i(S_i+1) - S(S+1)]^2 c_i^2}{\sum_{i=1}^K c_i^2}, \end{aligned} \tag{23}$$

where we used Eqs. (18), (19), and (21) and the orthonormality of the states  $\tilde{\zeta}_i$ .

So far, we have given formulas to compute and to eliminate the spin contamination of a trial wave function. In Appendix C, we illustrate the above formulas using the Li and Be atoms as examples and, in the case of Li, we prove that a spin-dependent two-body Jastrow factor always yields a non-zero spin contamination. The explicit formulas for computing the spin contamination in QMC are provided in Sec. V.

### V. EXPLICIT FORMULAS FOR QMC CALCULATIONS

We now show how the spin contamination is calculated in QMC calculations. Note that, from the exchange property of  $F_i$ ,

$$\int d^{3N} \mathbf{r} \sum_{i=1}^K F_i^2 = K \int d^{3N} \mathbf{r} F_1^2. \tag{24}$$

Therefore, we can sample  $F_1^2$  and compute  $\delta S^2$  using

$$\begin{aligned} \delta S^2 &= \int d^{3N} \mathbf{r} \frac{F_1^2}{\int d^{3N} \mathbf{r} F_1^2} \\ &\quad \times \left[ \frac{\sum_{i, S_i \neq S} [S_i(S_i+1) - S(S+1)]^2 \tilde{F}_i^2}{K \times F_1^2} \right]. \end{aligned} \tag{25}$$

The Monte Carlo implementation of this formula consists of sampling the electron coordinates from the normalized function

$$\frac{F_1^2}{\int d^{3N} \mathbf{r} F_1^2}, \tag{26}$$

and averaging the term in square brackets over the sampled points.

TABLE I. Spin contamination ( $\delta S^2$ ) obtained in VMC for different wave functions for the ground state of Li and Be.  $b$  and  $b'$  are parameters that enter in the Jastrow factor [Eq. (D6)]. The subscripts 1 and 0 indicate parallel and antiparallel, respectively.  $E_{\text{VMC}}$  is the total energy in VMC and  $\sigma_{\text{VMC}}$  is the root mean square fluctuation of the local energy in VMC. The exact nonrelativistic energy of the ground state is  $-7.478\,06$  hartrees for Li and  $-14.667\,36$  hartrees for Be (Ref. 10). All the energies are in hartree atomic units.

Atom	Wave function	$b_0$	$b_1$	$b'$	$E_{\text{VMC}}$	$\sigma_{\text{VMC}}$	$\delta S^2$
Li	$\Psi_{2\text{-body}}$	1/2	1/4	$b'_1 = b'_0$	-7.473 89(4)	0.243	0.000 097(2)
		1/2	1/4	$b'_1 \neq b'_0$	-7.474 27(4)	0.242	0.000 045(1)
	$\Psi_{3\text{-body}}$	1/2	1/2	$b'_1 = b'_0$	-7.477 80(1)	0.046	0.
		1/2	1/4	$b'_1 = b'_0$	-7.477 88(1)	0.038	0.000 050(1)
		1/2	1/4	$b'_1 \neq b'_0$	-7.477 88(1)	0.037	0.000 032(1)
	$\Psi_{4\text{-body}}$	1/2	1/4	$b'_1 = b'_0$	-7.477 96(1)	0.028	0.000 022(1)
		1/2	1/4	$b'_1 \neq b'_0$	-7.477 96(1)	0.028	0.000 009(0)
	Be	$\Psi_{2\text{-body}}$	1/2	1/4	$b'_1 = b'_0$	-14.660 71(5)	0.348
1/2			1/4	$b'_1 \neq b'_0$	-14.660 88(5)	0.345	0.000 13(1)
$\Psi_{3\text{-body}}$		1/2	1/2	$b'_1 = b'_0$	-14.665 45(1)	0.122	0.
		1/2	1/4	$b'_1 = b'_0$	-14.666 01(1)	0.095	0.000 31(1)
		1/2	1/4	$b'_1 \neq b'_0$	-14.666 62(1)	0.089	0.000 13(1)
$\Psi_{4\text{-body}}$		1/2	1/4	$b'_1 = b'_0$	-14.666 77(1)	0.076	0.000 23(2)
		1/2	1/4	$b'_1 \neq b'_0$	-14.666 77(1)	0.073	0.000 06(0)

The density operator is defined as  $\hat{\rho}(\mathbf{r}) = \sum_{i=1}^N \delta(\mathbf{r} - \mathbf{r}_i)$ , and the single-particle density for a normalized wave function  $\Psi$  is  $\sum_{s_1, \dots, s_N} \int d^3N \mathbf{r} \hat{\rho}(\mathbf{r}) \Psi^2$ . If  $\Psi$  is not an eigenstate of  $\hat{S}^2$ , the density associated with  $\Psi$  contains contributions from other spin components.

Suppose that the spin  $S$  state  $\Psi_S$  is the dominant component of  $\Psi$ . It can be projected out from  $\Psi$  using Eq. (18), and the portion of the density of  $\Psi$  resulting from other spin components is

$$\begin{aligned} \delta\rho(\mathbf{r}) &= \sum_{s_1, \dots, s_N} \int d^3N \mathbf{r} \hat{\rho}(\mathbf{r}) \frac{(\Psi^2 - \Psi_S^2)}{\sum_{s_1, \dots, s_N} \int d^3N \mathbf{r} \Psi^2} \\ &= \sum_{s_1, \dots, s_N} \int d^3N \mathbf{r} \frac{F_1^2}{\int d^3N \mathbf{r} F_1^2} \left[ \frac{\hat{\rho}(\mathbf{r})(\Psi^2 - \Psi_S^2)}{K \times F_1^2} \right] \\ &= \int d^3N \mathbf{r} \frac{F_1^2}{\int d^3N \mathbf{r} F_1^2} \left[ \frac{\hat{\rho}(\mathbf{r}) \sum_{i, s_i \neq S} \tilde{F}_i^2}{K \times F_1^2} \right]. \end{aligned} \quad (27)$$

Therefore, we sample the electron coordinates from the normalized function of Eq. (26) and average the term in square brackets of Eq. (27) over the sampled points.

Integrating Eq. (27), we obtain

$$\int d\mathbf{r} \delta\rho(\mathbf{r}) = N \frac{\sum_{i, s_i \neq S} c_i^2}{\sum_{i=1}^K c_i^2}. \quad (28)$$

Comparing Eqs. (23) and Eq. (28), it is clear that the numerical value of spin contamination can be used to estimate its integrated effect on the density. In the case that the spin contamination arises from a single spin component there is a precise relationship between the two quantities.

## VI. RESULT AND DISCUSSION

Using Eq. (25), we have calculated the spin contamination of several trial wave functions of the Li and Be atoms discussed in Appendix D. All the parameters occurring in the determinantal and Jastrow parts of the wave functions were optimized using the variance minimization method.<sup>5,6</sup>

For each of the two atoms, we construct seven wave functions. They are grouped according to the Jastrow factor used, i.e.,  $\Psi_{2\text{-body}}$ ,  $\Psi_{3\text{-body}}$ , and  $\Psi_{4\text{-body}}$ . The first wave function for  $\Psi_{3\text{-body}}$  has a completely symmetric Jastrow under interchange of the spatial coordinates of two electrons: the parallel-spin electrons are correlated with the same  $B$  term [Eq. (D6)] as the antiparallel ones ( $b_1 = b_0 = 1/2$  and  $b'_1 = b'_0$ ), so that the cusp condition for parallel spin electrons is violated. This wave function is an eigenstate of  $\hat{S}^2$ , i.e., the spin contamination is zero. The other six wave functions satisfy the cusp conditions both for antiparallel- and parallel-spin electrons ( $b_0 = 1/2$  and  $b_1 = 1/4$ ) but are not eigenstates of  $\hat{S}^2$ , i.e., the spin contamination is not zero. Each  $\Psi_{n\text{-body}}$  group contains two wave functions differing only in whether the constraint  $b'_1 = b'_0$  is imposed or  $b'_1$  and  $b'_0$  are allowed to vary independently.

In Table I, we list the spin contamination  $\delta S^2$  for our wave functions. We also give two quantities that are measures of the quality of the wave functions: the total energy obtained in variational Monte Carlo (VMC),  $E_{\text{VMC}}$ , and the root mean square fluctuations of the local energy in VMC,  $\sigma_{\text{VMC}}$ . For all the wave functions studied, the spin contamination is small.

The first thing to notice in Table I is that when higher correlations are included, not only are  $E_{\text{VMC}}$  and  $\sigma_{\text{VMC}}$  improved,<sup>4</sup> but the spin contamination is also significantly reduced.

The wave functions with the symmetric Jastrow factors,

which have no spin contamination ( $\delta S^2=0$ ) and violate the parallel-spin electron cusp conditions, yield values of  $\sigma_{\text{VMC}}$  that are larger by 21% and 24% for Li and 28% and 37% for Be, as compared to the other two three-body correlation wave functions which satisfy the cusp conditions.

The wave functions in which  $b'_1$  and  $b'_0$  are allowed to vary independently have spin contaminations that are much smaller than those for the unconstrained wave functions. In Li, spin contamination decreases by a factor of 2.2 for  $\Psi_{2\text{-body}}$ , 1.6 for  $\Psi_{3\text{-body}}$ , and 2.4 for  $\Psi_{4\text{-body}}$ , while, in Be, the corresponding reductions are 5.8, 2.4, and 3.8, respectively. This improvement occurs despite the fact that the improvements in the other two measures of wave function quality,  $E_{\text{VMC}}$  and  $\sigma_{\text{VMC}}$ , are relatively modest.

Quantum Monte Carlo has been used to calculate accurate charge densities<sup>11</sup> from which other accurate density functional quantities can also be obtained and compared with the corresponding quantities obtained from the commonly used approximate density functionals. Since the approximate functionals yield densities that differ only on the order of 1% or less from the true density in the region where the density is appreciable, in order for the QMC density to yield useful information, it should have an error that is no larger than say 0.1%. Hence, it is worth checking that the error in the QMC density arising from spin contamination is smaller than this. Since both Li and Be each have only one spurious spin component, from Eqs. (23) and (28), we obtain that the fraction of the density arising from the spurious spin component, integrated over space and divided by the total electronic charge  $N$ , is

$$\frac{\int d\mathbf{r} \delta\rho(\mathbf{r})}{N} = C \delta S^2, \quad (29)$$

where  $C=1/9$  for Li and  $C=1/36$  for Be. The wave functions used are similar to the third three-body wave functions for Li and Be in Table I which have spin contaminations of  $3.2 \times 10^{-5}$  and  $1.3 \times 10^{-4}$ , respectively. From Eq. (29) we obtain that  $\int d\mathbf{r} \delta\rho(\mathbf{r})/N = 3 \times 10^{-6}$  both for Li and Be which is negligible compared both to the required accuracy and to the largest source of error, viz. the use of the fixed-node approximation.

In conclusion, we have considered three options for constructing QMC wave functions. The first option, which for most purposes offers the best compromise, is to employ a spin-assigned wave function  $F_1$  with a Jastrow factor that satisfies the cusp conditions. This wave function will in general not be an eigenstate of  $\hat{S}^2$  but we have shown that such wave functions, with parameters optimized by the variance minimization technique, have very small spin contamination, and that including higher-order correlations in the Jastrow factor further reduces the spin contamination. A second possibility is to employ a totally symmetric Jastrow factor, but this results in a somewhat less accurate total energy and a large (20%–40%) increase in the root mean square fluctuations of the local energy, and is therefore not a good choice. A third possibility is to construct a wave function that both satisfies the cusp conditions and is an eigenstate of  $\hat{S}^2$ . As discussed in Appendix A, this increases the computational

cost in evaluating expectation values by a factor of  $N$ . For small systems, this choice is feasible and is probably the best choice if very high accuracy is required but it becomes impractical for large systems. A compromise solution is to construct wave functions of the first type by minimizing not only the variance of the local energies but also the spin contamination. This, of course, results in an increased computational cost of  $\mathcal{O}(N)$  for the optimization of the wave function but not for the Monte Carlo runs subsequently performed with the optimized wave functions. At the present time, it is not clear how much of a further reduction in the spin contamination could be achieved by doing this.

## ACKNOWLEDGMENTS

We thank John Morgan and Kevin Schmidt for helpful discussions, and David Ceperley for pointing out that the additional computational cost of computing expectation values of spin-projected wave functions increases only linearly with the number of electrons. The calculations were performed on the IBM SP2 computer at the Cornell Theory Center. This work is supported by the Office of Naval Research.

## APPENDIX A: EXPECTATION VALUES FOR SPIN-PROJECTED WAVE FUNCTIONS

In this appendix we show that, although each spin-projected wave function,  $\Psi_S$ , is comprised of  $\mathcal{O}(N!)$  spin-assigned wave functions  $F_i$ , the computational cost of evaluating expectation values of spin-independent operators is only  $\mathcal{O}(N)$  greater for  $\Psi_S$  than for a particular  $F_i$ , say  $F_1$ .

Following the same arguments used to derive Eq. (19), we obtain that the expectation value of a spin-independent operator,  $\mathcal{O}$ , can be expressed as,

$$\begin{aligned} \frac{\langle \psi_S | \mathcal{O} | \psi_S \rangle}{\langle \psi_S | \psi_S \rangle} &= \frac{\sum_{i=1}^{N_S} \int d^3N \mathbf{r} \bar{F}_i \mathcal{O} \bar{F}_i}{\sum_{i=1}^{N_S} \int d^3N \bar{F}_i^2} \\ &= \frac{\sum_{j=1}^K \sum_{k=1}^K (\sum_{i=1}^{N_S} T_{ij} T_{ik}) \int d^3N \mathbf{r} F_j \mathcal{O} F_k}{\sum_{j=1}^K \sum_{k=1}^K (\sum_{i=1}^{N_S} T_{ij} T_{ik}) \int d^3N \mathbf{r} F_j F_k}. \end{aligned} \quad (A1)$$

Therefore, the computational cost of evaluating the expectation value of such an operator on the spin-projected wave function  $\Psi_S$  depends on the number of distinct integrals,  $\int F_j \mathcal{O} F_k$  and  $\int F_j F_k$ . Since the operator  $\mathcal{O}$  is spin-independent and fully symmetric with respect to interchange of particle indices, the number of distinct integrals in the numerator and denominator are the same. For simplicity of notation, we compute this number for the denominator.

As mentioned in Sec. II, the components  $F_i$  are all the same except for a relabeling of the particle indices and a change in sign for odd permutations. Therefore, we can just consider the following overlap integrals,

$$\int d^3N \mathbf{r} F_1(\mathbf{r}_1, \dots, \mathbf{r}_N) F_k(\mathbf{r}_1, \dots, \mathbf{r}_N). \quad (A2)$$

Since the  $F_i$  are antisymmetric under the interchange of like-spin electrons, the above integrals depend (aside for a sign

change for odd permutations) only on the number of up-spin electrons that have been interchanged with down-spin electrons. This number can range from zero to  $\min(N_\uparrow, N_\downarrow)$ , resulting in  $\{\min(N_\uparrow, N_\downarrow) + 1\}$  distinct integrals, a number that is far smaller than the number of  $F_i$  that comprise  $\Psi_S$ , viz.  $N!/(N_\uparrow!N_\downarrow!)$ .

For a spin-unpolarized system, this number equals  $\{N/2 + 1\}$  but a further reduction is possible by exploiting the symmetry that  $F_1$  is invariant under the interchange of all up-spin electrons with all down-spin electrons,

$$F_1(\mathbf{r}_1, \dots, \mathbf{r}_{N/2}, \mathbf{r}_{N/2+1}, \dots, \mathbf{r}_N) = F_1(\mathbf{r}_{N/2+1}, \dots, \mathbf{r}_N, \mathbf{r}_1, \dots, \mathbf{r}_{N/2}). \quad (\text{A3})$$

The overlap integral obtained permuting  $\{[N/4] + j\}$  up-spin with as many down-spin electrons in  $F_1$  has the same value as the one obtained by permuting  $\{[N/4] - j\}$  unlike-spin electrons. Therefore, for a spin-unpolarized system, the number of distinct overlap integrals is given by  $\{[N/4] + 1\}$ .

## APPENDIX B: MORE ABOUT ELECTRON-ELECTRON CUSP CONDITIONS

For a given pair of electrons, the value of the cusp, according to Eq. (10), depends on the behavior of the wave function at the coincidence point. It is clear from the form of Eq. (7) that  $F_1$  is antisymmetric under the exchange of two parallel-spin electrons and consequently must vanish linearly when they coincide. Hence,  $l_0 = 1$  except possibly on a set of points of measure zero where  $l_0 \geq 3$ . For antiparallel-spin electrons, three values of  $l_0$  are possible,  $l_0 = 0, 1, 2$ . For most states,  $F_1$  is not either symmetric or antisymmetric under the exchange of two antiparallel-spin electrons and need not vanish as they approach each other. However, for certain states, it can be shown that  $F_1$  must be antisymmetric and for other states that it must be symmetric under the exchange of two antiparallel-spin electrons. Some of the  $d_i$  appearing in Eq. (5) may be related by symmetry. Symmetry related terms in the sum in Eq. (5) comprise configuration state functions (CSFs). If even one of the configuration state functions in Eq. (5) consists of a single determinant then  $l_0 = 0$  for that wave function. If all the CSFs consist of more than a single determinant then it is possible for the determinants within a CSF to cancel at the collision point leading to  $l_0 = 1$  or  $l_0 = 2$ . The latter possibility ( $l_0 = 2$ ) was pointed out by Morgan and Kutzelnigg<sup>12</sup> in connection with the ordering of levels in a multiplet and violations of Hund's rules. Of course, the usual prescription of using a Jastrow factor with a cusp of 1/2 for antiparallel-spin electrons fails if either  $l_0 = 1$  or  $l_0 = 2$ , the proper cusps being 1/4 and 1/6, respectively, in conformity with Eq. (10). States with  $l_0 = 1$  may be avoided by always choosing, for a given value of  $S$ , a state with extremal  $S_z$ , i.e.,  $S_z = \pm S$ .

### 1. $l_0 = 1$ case for antiparallel-spin electrons

We first give an example of a state where  $l_0 = 1$  at the coincidence of two antiparallel-spin electrons. Consider the  $S_z = 0$ ,  $^3S$  state resulting from a  $(1s)(2s)$  configuration as

occurs, for example, for the first excited state of a helium atom. The wave function is a combination of two determinants,

$$\Psi = \{(1s^+, 2s^-) - (1s^-, 2s^+)\}, \quad (\text{B1})$$

where  $(\phi_1, \phi_2)$  denotes a Slater determinant constructed from the orbitals  $\phi_1$  and  $\phi_2$ , + and - indicate up- and down-spin electrons, respectively, and an orbital is characterized by the quantum numbers,  $n$ ,  $l$ , and  $m_l$ . This expression can be rewritten as

$$\Psi = (1s(r_1)2s(r_2) - 2s(r_1)1s(r_2))\chi_t(s_1, s_2), \quad (\text{B2})$$

where  $\chi_t(s_1, s_2) = \chi_\uparrow(s_1)\chi_\downarrow(s_2) + \chi_\downarrow(s_1)\chi_\uparrow(s_2)$ . Hence, in Eq. (4),

$$F_1 = 1s(r_1)2s(r_2) - 2s(r_1)1s(r_2), \quad (\text{B3})$$

and in Eq. (7),  $D_1^+(\mathbf{r}_1) = 1s(r_1)$ ,  $D_1^-(\mathbf{r}_2) = 2s(r_2)$ ,  $D_2^+(\mathbf{r}_1) = 2s(r_1)$ , and  $D_2^-(\mathbf{r}_2) = 1s(r_2)$ . Note that, for the  $S_z = 0$  state, the simplest wave function with the correct symmetry has two determinants [as is clear from Eq. (B1)] whereas, for the  $S_z = 1$  and  $S_z = -1$  states, a single determinant suffices. This is obscured by the fact that  $F_1$  is identically the same for all three  $S_z$  states. However, from the form of the full antisymmetric wave function  $\Psi$  it is clear that  $F_1$  should be viewed as a single  $2 \times 2$  determinant for the  $S_z = \pm 1$  states but as the sum of products of  $1 \times 1$  determinants for the  $S_z = 0$  state. It is clear from Eq. (B3) that  $F_1$  is antisymmetric and  $l_0 = 1$ , and consequently the Jastrow factor must have a cusp of 1/4 rather than 1/2. We can avoid all occurrences of  $l_0 = 1$ , for antiparallel-spin electrons, if we calculate only states with  $|S_z| = S$ . Since the expectation values of all spin-independent operators are independent of which  $S_z$  state we choose, this is not a restriction.

### 2. $l_0 = 2$ case for antiparallel-spin electrons

Next, following Ref. 12, we give an example of a state where  $l_0 = 2$ . We consider the  $^1P$  state of two electrons in a  $(2p)(3p)$  parent configuration. The wave function is the combination of four determinants,

$$\Psi = \{(2p_1^+, 3p_0^-) - (2p_1^-, 3p_0^+)\} - \{(2p_0^+, 3p_1^-) - (2p_0^-, 3p_1^+)\} \quad (\text{B4})$$

$$= (2p_1(\mathbf{r}_1)3p_0(\mathbf{r}_2) + 3p_0(\mathbf{r}_1)2p_1(\mathbf{r}_2))\chi_s(s_1, s_2) - (2p_0(\mathbf{r}_1)3p_1(\mathbf{r}_2) + 3p_1(\mathbf{r}_1)2p_0(\mathbf{r}_2)) \times \chi_s(s_1, s_2) \quad (\text{B5})$$

$$= [R_{21}(r_1)R_{31}(r_2) - R_{31}(r_1)R_{21}(r_2)] \times [Y_{11}(\Omega_1)Y_{10}(\Omega_2) - Y_{10}(\Omega_1)Y_{11}(\Omega_2)] \times \chi_s(s_1, s_2), \quad (\text{B6})$$

where  $\chi_s(s_1, s_2) = \chi_\uparrow(s_1)\chi_\downarrow(s_2) - \chi_\downarrow(s_1)\chi_\uparrow(s_2)$ . Hence,  $F_1$  vanishes quadratically at the coincidence point since each of the factors in square brackets vanishes linearly.

Unlike the  $l_0 = 1$  case, the  $l_0 = 2$  case cannot be circumvented by considering a different state with the same expectation values. However, it occurs only rarely for ground states<sup>12</sup> and consequently the usual prescription of using a

Jastrow with a cusp of 1/4 (1/2) for parallel- (antiparallel)-spin electrons is correct for most ground state wave functions.

**APPENDIX C: ILLUSTRATION USING SIMPLE EXAMPLES**

**1. Li ground state**

The Li ground state has symmetry  $^2S$ . For  $S_z = 1/2$ , there are three independent spin functions,

$$\begin{aligned} \zeta_1 &= \chi_{\downarrow}(s_1)\chi_{\uparrow}(s_2)\chi_{\uparrow}(s_3), \\ \zeta_2 &= \chi_{\uparrow}(s_1)\chi_{\downarrow}(s_2)\chi_{\uparrow}(s_3), \\ \zeta_3 &= \chi_{\uparrow}(s_1)\chi_{\uparrow}(s_2)\chi_{\downarrow}(s_3), \end{aligned} \tag{C1}$$

and the total space-spin wave function of Li can be written as

$$\Psi = \sum_{i=1}^3 F_i(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)\zeta_i(s_1, s_2, s_3). \tag{C2}$$

Antisymmetry of the wave function implies that

$$\begin{aligned} F_1(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) &= -F_1(\mathbf{r}_1, \mathbf{r}_3, \mathbf{r}_2) = F_2(\mathbf{r}_3, \mathbf{r}_1, \mathbf{r}_2) \\ &= -F_2(\mathbf{r}_2, \mathbf{r}_1, \mathbf{r}_3) = F_3(\mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_1) \\ &= -F_3(\mathbf{r}_3, \mathbf{r}_2, \mathbf{r}_1). \end{aligned} \tag{C3}$$

Since  $S = S_z = 1/2$ ,  $S_z^2 - S_z + N_{\uparrow} = 7/4$ , the diagonal terms of  $\Theta$  are equal to 7/4 and nondiagonal terms are equal to 1, so that  $(\Theta_{ij} - S(S+1)\delta_{ij}) = 1$ . The orthonormalized eigenvectors of  $\Theta$  are

$$\begin{aligned} t_1 &= \frac{1}{\sqrt{2}}(1, 0, -1), & t_2 &= \frac{1}{\sqrt{6}}(1, -2, 1), \\ t_3 &= \frac{1}{\sqrt{3}}(1, 1, 1), \end{aligned} \tag{C4}$$

with eigenvalues 3/4, 3/4, 15/4, corresponding to total spin 1/2, 1/2, 3/2, respectively.

The normalized wave function,  $\Psi$ , can be decomposed, according to Eqs. (11) and (13), into

$$\Psi = \Psi_{1/2} + \Psi_{3/2} = c_{1/2}\psi_{1/2} + c'_{1/2}\psi'_{1/2} + c_{3/2}\psi_{3/2}. \tag{C5}$$

Using Eqs. (17) and (18), we can express  $\Psi_{1/2}$  and  $\Psi_{3/2}$  in terms of  $F_i$  and  $\zeta_i$ ,

$$\begin{aligned} \Psi_{1/2} &= \tilde{F}_1 \tilde{\zeta}_1 + \tilde{F}_2 \tilde{\zeta}_2 \\ &= \frac{(F_1 - F_3)(\zeta_1 - \zeta_3)}{2} \\ &\quad + \frac{(F_1 - 2F_2 + F_3)(\zeta_1 - 2\zeta_2 + \zeta_3)}{6} \\ &= \frac{1}{3}[(2F_1 - F_2 - F_3)\zeta_1 + (2F_2 - F_1 - F_3)\zeta_2 \\ &\quad + (2F_3 - F_1 - F_2)\zeta_3], \end{aligned} \tag{C6}$$

$$\Psi_{3/2} = \tilde{F}_3 \tilde{\zeta}_3 = \frac{1}{3} \left( \sum_{i=1}^3 F_i \right) \left( \sum_{j=1}^3 \zeta_j \right). \tag{C7}$$

Therefore, from Eq. (C7),  $\Psi$  is an eigenstate of  $\hat{S}^2$  with  $S = 1/2$  if  $\Psi_{3/2} = 0$  or  $\sum_{i=1}^3 F_i = 0$ . Substituting this in Eq. (C7) we recover  $\Psi_{1/2} = \sum_{i=1}^3 F_i \zeta_i = \Psi$ . From Eq. (23) and  $\tilde{F}_3^2 = 1/3(F_1 + F_2 + F_3)^2$ , we obtain

$$\delta S^2 = 3 \frac{\int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 (\sum_{i=1}^3 F_i)^2}{\int d\mathbf{r}_1 d\mathbf{r}_2 d\mathbf{r}_3 \sum_{i=1}^3 F_i^2} = \frac{9c_{3/2}^2}{\sum_i c_i^2}. \tag{C8}$$

As discussed in Sec. IV, the  $\tilde{F}_i$ 's belonging to a degenerate subspace need not be antisymmetric but their sum is. Here we have an example of this. The  $\psi_{1/2}$  and  $\psi'_{1/2}$  are not antisymmetric but their sum  $\Psi_{1/2}$  is. Finally, note that each of  $\Psi_{1/2}$  and  $\Psi_{3/2}$  of Eqs. (C6) and (C7) are antisymmetric, satisfy the cusp conditions, and are eigenstates of  $\hat{S}^2$ .

We now prove, in the special case of a single-determinant wave function for Li, that the use of a two-body Jastrow factor that is not totally symmetric in the spatial coordinates must result in a spin-contaminated wave function. The necessary and sufficient condition for  $\Psi$  to be an  $S = 1/2$  eigenstate of  $\hat{S}^2$  is

$$\begin{aligned} \sum_{i=1}^3 F_i(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) &= F_1(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) + F_1(\mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_1) \\ &\quad + F_1(\mathbf{r}_3, \mathbf{r}_1, \mathbf{r}_2) = 0. \end{aligned} \tag{C9}$$

A simple choice for  $F_1$  is

$$\begin{aligned} F_1(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) &= 1s(r_1)[1s(r_2)2s(r_3) \\ &\quad - 2s(r_2)1s(r_3)], \end{aligned} \tag{C10}$$

where  $1s(r)$  and  $2s(r)$  are single-particle orbitals. Clearly this choice for  $F_1$  satisfies Eq. (C9) and the corresponding full wave function has definite angular momentum  $L = L_z = 0$ , spin eigenvalues,  $S = S_z = 1/2$ , and is completely antisymmetric.

Introducing a Jastrow factor, we rewrite  $F_1$  as

$$\begin{aligned} F_1(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) &= 1s(r_1)[1s(r_2)2s(r_3) \\ &\quad - 2s(r_2)1s(r_3)]J(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3). \end{aligned} \tag{C11}$$

Eq. (C9) is satisfied if the Jastrow factor is completely symmetric,

$$J(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) = J(\mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_1) = J(\mathbf{r}_3, \mathbf{r}_1, \mathbf{r}_2), \tag{C12}$$

but may not be satisfied otherwise. Substituting  $F_1$  of Eq. (C11) into Eq. (C9) with the particle indices properly permuted, we obtain

$$\begin{aligned} 1s(r_2)1s(r_3)2s(r_1)[J(\mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_1) - J(\mathbf{r}_3, \mathbf{r}_1, \mathbf{r}_2)] + 1s(r_3) \\ \times 1s(r_1)2s(r_2)[J(\mathbf{r}_3, \mathbf{r}_1, \mathbf{r}_2) - J(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3)] + 1s(r_1) \\ \times 1s(r_2)2s(r_3)[J(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3) - J(\mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_1)] = 0. \end{aligned} \tag{C13}$$

The three orbital terms  $1s(r_2)1s(r_3)2s(r_1)$ ,  $1s(r_3)1s(r_1)2s(r_2)$ , and  $1s(r_1)1s(r_2)2s(r_3)$  are linearly independent. Consequently, if  $J$  is a function of only  $r_{12}, r_{13}, r_{23}$  [as in Eq. (D1)], it follows from the algebraic independence of  $r_1, r_2, r_3, r_{12}, r_{13}, r_{23}$  that each of the terms in square parentheses must be zero. So, in order to satisfy Eq. (C13), the Jastrow factor  $J$  must be completely symmetric in the spatial coordinates as was pointed out in Ref. 1. In

general,  $J$  could also depend on the distances of the electrons from the nucleus,  $r_1, r_2, r_3$  [as in Eq. (D2)]. In that case the terms in the square brackets are not independent of the orbital terms and the proof no longer holds.

## 2. Be ground state

The ground state of Be is  $^1S$ . Since  $N=4$  and  $S_z=0$  ( $N_\uparrow=2, N_\downarrow=2$ ), there are  $4!/(2!2!)=6$  spin states  $\zeta_i$ ,

$$\begin{aligned}\zeta_1 &= \chi_\uparrow(s_1)\chi_\uparrow(s_2)\chi_\downarrow(s_3)\chi_\downarrow(s_4), \\ \zeta_2 &= \chi_\uparrow(s_1)\chi_\downarrow(s_2)\chi_\uparrow(s_3)\chi_\downarrow(s_4), \\ \zeta_3 &= \chi_\downarrow(s_1)\chi_\uparrow(s_2)\chi_\uparrow(s_3)\chi_\downarrow(s_4), \\ \zeta_4 &= \chi_\downarrow(s_1)\chi_\downarrow(s_2)\chi_\uparrow(s_3)\chi_\uparrow(s_4), \\ \zeta_5 &= \chi_\downarrow(s_1)\chi_\uparrow(s_2)\chi_\downarrow(s_3)\chi_\uparrow(s_4), \\ \zeta_6 &= \chi_\uparrow(s_1)\chi_\downarrow(s_2)\chi_\downarrow(s_3)\chi_\uparrow(s_4).\end{aligned}\quad (\text{C14})$$

The  $\Theta$  matrix is given by

$$\Theta = \begin{bmatrix} 211011 \\ 121101 \\ 112110 \\ 011211 \\ 101121 \\ 110112 \end{bmatrix}.\quad (\text{C15})$$

The orthonormalized eigenvectors of  $\Theta$  are

$$\begin{aligned}t_1 &= \frac{1}{\sqrt{12}}(1,1,-2,1,1,-2), & t_2 &= \frac{1}{2}(-1,1,0,-1,1,0), \\ t_3 &= \frac{1}{\sqrt{2}}(0,0,-1,0,0,1), \\ t_4 &= \frac{1}{\sqrt{2}}(0,-1,0,0,1,0), & t_5 &= \frac{1}{\sqrt{2}}(-1,0,0,1,0,0), \\ t_6 &= \frac{1}{\sqrt{6}}(1,1,1,1,1,1),\end{aligned}\quad (\text{C16})$$

with eigenvalues 0,0,2,2,6 corresponding to total spin  $S=0,0,1,1,2$ , respectively.

The normalized approximate wave function can be decomposed according to Eqs. (11) and (13),

$$\begin{aligned}\Psi &= \Psi_0 + \Psi_1 + \Psi_2 = c_0\psi_0 + c'_0\psi'_0 + c_1\psi_1 + c'_1\psi'_1 + c''_1\psi''_1 \\ &\quad + c_2\psi_2.\end{aligned}\quad (\text{C17})$$

The  $\tilde{F}_i$  defined in Eq. (17) are

$$\begin{aligned}\tilde{F}_1 &= \frac{1}{\sqrt{12}}[(F_1 + F_2 + F_4 + F_5) - 2(F_3 + F_6)], \\ \tilde{F}_2 &= \frac{1}{2}[(F_2 + F_5) - (F_1 + F_4)], \\ \tilde{F}_3 &= \frac{1}{\sqrt{2}}(F_6 - F_3),\end{aligned}\quad (\text{C18})$$

$$\tilde{F}_4 = \frac{1}{\sqrt{2}}(F_5 - F_2),$$

$$\tilde{F}_5 = \frac{1}{\sqrt{2}}(F_4 - F_1),$$

$$\tilde{F}_6 = \frac{1}{\sqrt{6}} \sum_{i=1}^6 F_i.$$

Antisymmetry of the wave function implies that

$$F_1(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) = F_4(\mathbf{r}_3, \mathbf{r}_4, \mathbf{r}_1, \mathbf{r}_2).\quad (\text{C19})$$

The trial wave functions commonly used have the symmetry

$$F_1(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) = F_1(\mathbf{r}_3, \mathbf{r}_4, \mathbf{r}_1, \mathbf{r}_2).\quad (\text{C20})$$

Consequently,

$$F_1(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4) = F_4(\mathbf{r}_1, \mathbf{r}_2, \mathbf{r}_3, \mathbf{r}_4)\quad (\text{C21})$$

and  $\tilde{F}_5=0$ . Similarly,  $\tilde{F}_3=\tilde{F}_4=0$ . Since  $\tilde{F}_3, \tilde{F}_4, \tilde{F}_5$  are the  $S=1$  component of  $\Psi$ ,  $c_1=c'_1=c''_1=0$ . So, if the form of the trial wave function is such that Eq. (C20) is satisfied, the spin contamination comes purely from the  $S=2$  component and, from Eqs. (23) and (C18), we obtain,

$$\delta S^2 = 6 \frac{\int d\mathbf{r}_1 \dots d\mathbf{r}_4 (\sum_{i=1}^6 F_i)^2}{\int d\mathbf{r}_1 \dots d\mathbf{r}_4 \sum_{i=1}^6 F_i^2} = \frac{36c_2^2}{\sum_i c_i^2}.\quad (\text{C22})$$

## APPENDIX D: FORM OF THE WAVE FUNCTIONS USED

All the wave functions we tested have the form of Eq. (7) but differ in the body order of the Jastrow factor. The two-body Jastrow factor is,

$$J_{2\text{-body}} = \prod_{i<j} \exp\left(\frac{b_{|s_{zi}+s_{zj}|} r_{ij}}{1 + b'_{s_{zi}+s_{zj}} r_{ij}}\right),\quad (\text{D1})$$

where the product is over all pairs of electrons. As discussed in Appendix B, the requirement that the local energy does not diverge at the electron-electron coincidence points results in cusp conditions on the wave function and fixes the value of  $b$  for antiparallel- and parallel-spin electrons to be  $b_0=1/2$  or  $1/6$  (choosing  $S_z=\pm S$  states) and  $b_1=1/4$ . For  $S=0$  Be ground state  $b'_{-1}=b'_1$ , and for Li there is only one spin-down electron, so we limit ourselves to  $b'_{-1}=b'_1$  but we consider both wave functions where the constraint  $b'_1=b'_0$  is imposed and wave functions where  $b'_1$  is allowed to vary independently of  $b'_0$ .

It is known that a generalized Jastrow factor that explicitly correlates two electrons and a nucleus yields significantly better energies and smaller variance of the local

energies.<sup>6</sup> A similar functional form was also used by Filippi and Umrigar<sup>1</sup> for first-row molecules where the generalized Jastrow factor is written as a product of terms of increasing body-order correlation, up to three-body (two electrons and a nucleus) correlation. The wave function includes terms to describe the electron–nucleus correlation,  $A$ , the electron–electron correlation,  $B_{s_{zi}+s_{zj}}$ , and the electron–electron–nucleus correlation,  $C$ ,

$$J_{3\text{-body}} = \prod_{\alpha i} \exp(A(r_{i\alpha})) \prod_{ij} \exp(B_{s_{zi}+s_{zj}}(r_{ij})) \\ \times \prod_{\alpha ij} \exp(C(r_{ij}, r_{i\alpha}, r_{j\alpha})). \quad (\text{D2})$$

The generalized Jastrow is expressed in terms of scaled interparticle distances,

$$R_{i\alpha} = (1 - e^{-\kappa r_{i\alpha}})/\kappa, \quad R_{ij} = (1 - e^{-\kappa r_{ij}})/\kappa \quad (\text{D3})$$

and

$$A = \frac{aR_{i\alpha}}{1 + a'R_{i\alpha}}, \quad (\text{D4})$$

$$B_{s_{zi}+s_{zj}} = \frac{b|s_{zi}+s_{zj}|R_{ij}}{1 + b'|s_{zi}+s_{zj}|R_{ij}}, \quad (\text{D5})$$

$$C = P_C(R_{i\alpha}, R_{j\alpha}, R_{ij}) + F(R_{i\alpha}, R_{j\alpha}, R_{ij}), \quad (\text{D6})$$

where  $P_C$  is a complete 5th order polynomial in  $R_{i\alpha}, R_{j\alpha}, R_{ij}$ , and  $F$  consists of the Fock-expansion motivated terms.<sup>13</sup> Note that the functions  $A$  and  $C$  in Eq. (D2) are spin independent but that  $B$  is allowed to be spin dependent in order to satisfy the cusp conditions.

Higher body correlations (three electrons and a nucleus) are utilized in the Jastrow factor of Huang, Umrigar, and Nightingale,<sup>4</sup> where invariant theory is applied to design an efficient algorithm to compute high order correlations. Although the generalized Jastrow factor that they use is not explicitly factorized into terms of increasing body-order correlation (invariants are used instead), nevertheless, the form of the wave function is equivalent to

$$J_{4\text{-body}} = J_{3\text{-body}}(r_{i\alpha}, r_{j\alpha}, r_{ij}) \\ \times \prod_{\alpha ijk} \exp(D(r_{ij}, r_{ik}, r_{jk}, r_{i\alpha}, r_{j\alpha}, r_{k\alpha})), \quad (\text{D7})$$

$$D = P_D(R_{i\alpha}, R_{j\alpha}, R_{k\alpha}, R_{ij}, R_{ik}, R_{jk}), \quad (\text{D8})$$

where  $P_D$  is a complete 5th order polynomial in the scaled variables and it only includes terms describing the electron–electron–electron and the electron–electron–electron–nucleus correlations.

For a detailed description of these wave functional forms, see Refs. 1 and 4 and references therein.

In the present paper, the parameters in the two-body Jastrow of Eq. (D1) were chosen as follows. The parameter  $b_0$  was always chosen to be 1/2 in order to satisfy the antiparallel-spin cusp condition. Wave functions with  $b_1 = 1/2$  and  $b_1 = 1/4$  were tested. The former yields a wave function with no spin contamination while the latter yields a wave function that obeys the cusp conditions. For the  $b_1 = 1/2$ , we only considered  $b'_1 = b'_0$  in order to have no spin contamination. For the  $b_1 = 1/4$ , we considered both wave functions for which  $b'_1 = b'_0$ , and wave functions where  $b'_1$  was allowed to vary independently of  $b'_0$ . As discussed in Ref. 1, a better constrained choice of  $b'_1$  is obtained by imposing that the logarithmic derivatives of the  $B$  terms in the Jastrow factor [Eq. (C2)] are equal for parallel and antiparallel spin electrons at large interelectron distances. This yields different values of  $b'_0$  and  $b'_1$ , respectively, that are related through the scale factor  $\kappa$  [cf. Eq. (D3)] as

$$b'_1 = \kappa \left[ \frac{1}{\sqrt{2}} \left( 1 + \frac{b'_0}{\kappa} \right) - 1 \right]. \quad (\text{D9})$$

We find in fact that, when we allow  $b'_1$  to vary independently of  $b'_0$ , its optimized value is close to that obtained from this formula.

<sup>1</sup>C. Filippi and C. J. Umrigar, *J. Chem. Phys.* **105**, 213 (1996).

<sup>2</sup>J. S. Andrews, D. Jayatilaka, R. G. A. Bone, N. C. Handy, and R. D. Amos, *Chem. Phys. Lett.* **183**, 423 (1991).

<sup>3</sup>J. Baker, A. Scheiner, and J. Andzelm, *Chem. Phys. Lett.* **216**, 380 (1993).

<sup>4</sup>C. Huang, C. J. Umrigar, and M. P. Nightingale, *J. Chem. Phys.* **107**, 3007 (1997).

<sup>5</sup>R. L. Coldwell, *Int. J. Quantum Chem., Symp.* **11**, 215 (1977).

<sup>6</sup>C. 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).

<sup>7</sup>T. Kato, *Commun. Pure Appl. Math.* **10**, 151 (1957).

<sup>8</sup>R. T. Pack and W. Byers-Brown, *J. Chem. Phys.* **45**, 556 (1966).

<sup>9</sup>P. O. Löwdin, *Phys. Rev.* **97**, 1509 (1955).

<sup>10</sup>E. R. Davidson, S. A. Hagstrom, S. J. Chakravorty, V. M. Umar, and C. F. Fischer, *Phys. Rev. A* **44**, 7071 (1991).

<sup>11</sup>C. J. Umrigar and X. Gonze, in *High Performance Computing and its Application to the Physical Sciences*, Proceedings of the Mardi Gras '93 Conference, edited by D. A. Browne *et al.* (World Scientific, Singapore, 1993); C. Filippi, C. J. Umrigar, and X. Gonze in *Theoretical and Computational Chemistry: Recent Developments in Density Functional Theory*, edited by J. Seminario (Elsevier, New York, 1996).

<sup>12</sup>J. D. Morgan III and W. Kutzelnigg, *J. Chem. Phys.* **97**, 2425 (1993).

<sup>13</sup>C. R. Myers, C. J. Umrigar, J. P. Sethna, and J. D. Morgan, *Phys. Rev. A* **44**, 5537 (1991).