

## Jastrow correlation factor for atoms, molecules, and solids

N. D. Drummond, M. D. Towler, and R. J. Needs

*TCM Group, Cavendish Laboratory, University of Cambridge, Madingley Road, Cambridge CB3 0HE, United Kingdom*

(Received 19 August 2004; revised manuscript received 13 October 2004; published 17 December 2004)

A form of Jastrow factor is introduced for use in quantum Monte Carlo simulations of finite and periodic systems. Test data are presented for atoms, molecules, and solids, including both all-electron and pseudopotential atoms. We demonstrate that our Jastrow factor is able to retrieve a large fraction of the correlation energy.

DOI: 10.1103/PhysRevB.70.235119

PACS number(s): 71.15.Dx, 31.25.-v

### I. INTRODUCTION

Many-electron wave functions may be accurately and compactly approximated by a product of a small number of Slater determinants and a positive Jastrow correlation factor. The Jastrow factor is an explicit function of the electron-electron separations, so that expectation values calculated with a Slater-Jastrow wave function do not separate in the electron coordinates. Nevertheless, the variational and diffusion quantum Monte Carlo (VMC and DMC) methods permit the use of such explicitly correlated wave functions.

In VMC, expectation values are calculated using an approximate trial wave function, the integrals being performed by a Monte Carlo method. In DMC (Refs. 1 and 2) the imaginary-time Schrödinger equation is used to evolve an ensemble of electronic configurations towards the ground state. The fermionic symmetry is maintained by the fixed-node approximation,<sup>3</sup> in which the nodal surface of the wave function is constrained to equal that of a trial wave function. The DMC method gives the energy that would be obtained in a VMC calculation with the same Slater determinants, but using the best possible Jastrow factor.

Although the DMC energy is in principle independent of the Jastrow factor, a poor trial wave function increases the statistical error bars and the time-step and population-control biases. When nonlocal pseudopotentials are used within DMC, the locality approximation<sup>4,5</sup> leads to additional errors which are second order in the error in the trial wave function.<sup>6</sup> The expectation values of operators that do not commute with the Hamiltonian are often evaluated using extrapolated estimation,<sup>2</sup> the accuracy of the extrapolation depending on the quality of the trial wave function. In practice the efficiency and accuracy of both VMC and DMC calculations are critically dependent on the availability of high-quality Jastrow factors.

Our Jastrow factor is designed for use in atoms, molecules, and solids. We have used it in a variety of systems, and here we report results on the He, Ne<sup>8+</sup>, Ne, and Ni atoms, the NiO and SiH<sub>4</sub> molecules, and crystalline Si in the diamond structure. These systems include all-electron and pseudopotential descriptions of atoms, with the total number of electrons varying from 2 to 216. We pay particular attention to the issue of cutting off terms in the Jastrow factor at finite ranges, which is desirable because of the local nature of the inhomogeneous correlations in many systems, as well as for reasons of computational efficiency in large systems.

We obtained the values of the free parameters in our Jastrow factors by minimizing the variance of the energy.<sup>7,8</sup> All of our QMC calculations were performed using the CASINO package.<sup>9</sup> We use Hartree atomic units  $\hbar = |e| = m_e = 4\pi\epsilon_0 = 1$  throughout this article.

The rest of this paper is organized as follows. In Sec. II we describe the general form of our Jastrow factor, while in Sec. III we show how the electron-electron and electron-nucleus cusp conditions<sup>10</sup> apply to this form. The behavior of the local energy at electron-electron and electron-nucleus coalescence points is discussed in Sec. IV. Section V describes the Jastrow factor in detail. In Sec. VI we make further comments on the form of our Jastrow factor, while in Sec. VII we define our notation for specifying the Jastrow factor and give our criterion for judging its quality. In Secs. VIII–XI we report the results of studies of various systems. Finally, we draw our conclusions in Sec. XII.

### II. GENERAL FORM OF THE JASTROW FACTOR

The Slater-Jastrow wave function can be written as

$$\Psi(\{\mathbf{r}_i\}, \{\mathbf{r}_J\}) = \exp[J(\{\mathbf{r}_i\}, \{\mathbf{r}_J\})]D(\{\mathbf{r}_i\}), \quad (1)$$

where  $\{\mathbf{r}_i\}$  and  $\{\mathbf{r}_J\}$  denote the electron and ion coordinates, respectively,  $\exp[J]$  is the Jastrow factor, and  $D$  denotes the Slater part, which depends only implicitly on the  $\{\mathbf{r}_J\}$ . An accurate and efficient Jastrow factor should possess a number of qualities. The functional form of the Jastrow factor should be chosen to reflect the physics of the correlations in the system, and it should be parametrized efficiently. The trial wave function must be continuous everywhere and its gradient must be continuous wherever the potential is finite, so that the kinetic energy is well defined. The Kato cusp conditions<sup>10</sup> determine the behavior of the many-body wave function when two electrons, or an electron and a nucleus, are coincident. The cusp conditions derive from the requirement that the divergence in the local kinetic energy at a coalescence point cancels the divergence in the local potential energy. Failure to satisfy the cusp conditions leads to divergences in the local energy  $\Psi^{-1}\hat{H}\Psi$ , where  $\hat{H}$  is the Hamiltonian. These divergences are especially harmful in DMC calculations, where they can lead to population-control problems and significant biases. It is standard practice to use the Jastrow factor to enforce the cusp conditions. The Slater

part of the wave function is chosen to satisfy the correct symmetry under exchange of electrons, and therefore the Jastrow factor should be symmetric under exchange. Indeed the Slater part of the wave function is normally chosen to have the correct symmetries of the state, so we should choose a Jastrow factor that does not change this symmetry. Finally, the Jastrow factor should allow rapid evaluation, as this is one of the more computationally demanding parts of VMC and DMC calculations.

Our Jastrow factor is the sum of homogeneous, isotropic electron-electron terms  $u$ , isotropic electron-nucleus terms  $\chi$  centered on the nuclei, isotropic electron-electron-nucleus terms  $f$ , also centered on the nuclei and, in periodic systems, plane-wave expansions of electron-electron separation and electron position  $p$  and  $q$ . The form is

$$\begin{aligned}
 J(\{\mathbf{r}_{ij}\}, \{\mathbf{r}_I\}) = & \sum_{i=1}^{N-1} \sum_{j=i+1}^N u(r_{ij}) + \sum_{I=1}^{N_{\text{ions}}} \sum_{i=1}^N \chi_I(r_{iI}) \\
 & + \sum_{I=1}^{N_{\text{ions}}} \sum_{i=1}^{N-1} \sum_{j=i+1}^N f_I(r_{iI}, r_{jI}, r_{ij}) + \sum_{i=1}^{N-1} \sum_{j=i+1}^N p(\mathbf{r}_{ij}) \\
 & + \sum_{i=1}^N q(\mathbf{r}_i), \quad (2)
 \end{aligned}$$

where  $N$  is the number of electrons,  $N_{\text{ions}}$  is the number of ions,  $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ , and  $\mathbf{r}_{iI} = \mathbf{r}_i - \mathbf{r}_I$ . Note that  $u$ ,  $\chi$ ,  $f$ ,  $p$ , and  $q$  may also depend on the spins of  $i$  and  $j$ . Although we will present results using spin-dependent parameters, for compactness the spin type has been suppressed in all formulas. The basic form is not novel, as terms of each type present in Eq. (2) have appeared in Jastrow factors in the literature,<sup>2</sup> but our particular forms of  $u$ ,  $\chi_I$ , and  $f_I$  are new.

The plane-wave term  $p$  will describe similar sorts of correlation to the  $u$  term. In periodic systems the  $u$  term must be cut off at a distance less than or equal to the Wigner-Seitz radius of the simulation cell (see Sec. V C and therefore the  $u$  function includes electron pairs over less than three quarters of the simulation cell. The  $p$  term adds variational freedom in the ‘‘corners’’ of the simulation cell, which could be important in small cells. The  $p$  term can also describe anisotropic correlations, such as might be encountered in a layered compound. However, we expect that the  $u$  term will be considerably more important than the  $p$  term, which cannot describe the electron-electron cusps and is therefore best limited to describing longer-ranged correlations. The  $q$  term will describe similar electron-nucleus correlations to the  $\chi_I$  terms.

### III. THE ELECTRON-ELECTRON AND ELECTRON-NUCLEUS CUSP CONDITIONS

Imposing the cusp conditions on the Jastrow factor is non-trivial because the variables  $r_{ij}$ ,  $r_{iI}$ , and  $r_{jI}$  are not independent. It is important to understand the meaning of the derivatives considered in this section. In Eq. (3), for example, the derivative  $\partial\hat{\Psi}/\partial r_{ij}$  means the derivative with respect to  $r_{ij}$  with all other coordinates held fixed, while in Eq. (5) the derivative  $\partial J/\partial r_{ij}$  means the derivative with respect to  $r_{ij}$  with  $r_i$  and  $r_j$  fixed.

#### A. The antiparallel-spin electron-electron cusp condition

Consider the situation where two electrons of opposite spin,  $i$  and  $j$ , approach one another and the wave function is nonzero at the two-particle coalescence point. This condition holds at almost all coalescence points of antiparallel-spin electrons. Let us omit the coordinates of all the other electrons and write the wave function in terms of the center-of-mass and difference coordinates of electrons  $i$  and  $j$ ,  $\bar{\mathbf{r}}_{ij} = (\mathbf{r}_i + \mathbf{r}_j)/2$  and  $\mathbf{r}_{ij} = \mathbf{r}_i - \mathbf{r}_j$ . The cusp condition<sup>10</sup> is

$$\left( \frac{\partial \hat{\Psi}}{\partial r_{ij}} \right)_{r_{ij}=0} = \frac{1}{2} \hat{\Psi}_{r_{ij}=0}, \quad (3)$$

where  $\hat{\Psi}(\bar{\mathbf{r}}_{ij}, r_{ij})$  is the spherical average of  $\Psi(\bar{\mathbf{r}}_{ij}, \mathbf{r}_{ij})$  about the coalescence point.

Neglecting the cusplike  $p$  and  $q$  terms, the Slater-Jastrow wave function may be written as

$$\Psi(\bar{\mathbf{r}}_{ij}, \mathbf{r}_{ij}) = \exp[J(r_i, r_j, r_{ij})] D(\bar{\mathbf{r}}_{ij}, \mathbf{r}_{ij}), \quad (4)$$

where for clarity we have assumed there is only one nucleus, which is located at the origin. Consider the change in the value of  $\Psi$  for a small displacement from the coalescence point such that the center-of-mass remains fixed:

$$\begin{aligned}
 \delta\Psi = \Psi_{r_{ij}=0} \times & \left\{ \left[ \left( \frac{\partial J}{\partial r_i} \right) - \left( \frac{\partial J}{\partial r_j} \right) \right]_{r_{ij}=0} \delta\mathbf{r}_i + \left( \frac{\partial J}{\partial r_{ij}} \right)_{r_{ij}=0} r_{ij} \right\} \\
 & + \exp[J_{r_{ij}=0}] (\nabla_{ij} D)_{r_{ij}=0} \cdot \mathbf{r}_{ij} + \mathcal{O}(r_{ij}^2), \quad (5)
 \end{aligned}$$

where  $\delta\mathbf{r}_i$  and  $\delta\mathbf{r}_j$  are the changes in  $\mathbf{r}_i$  and  $\mathbf{r}_j$  when the electron separation  $\mathbf{r}_{ij}$  is increased from zero, and we have used  $\delta\mathbf{r}_j = -\delta\mathbf{r}_i$ . If the spherical average about the coalescence point is taken then the terms involving  $\delta\mathbf{r}_i$  and  $\mathbf{r}_{ij}$  vanish to  $\mathcal{O}(r_{ij})$ , so that

$$\delta\hat{\Psi} = \hat{\Psi}_{r_{ij}=0} \left( \frac{\partial J}{\partial r_{ij}} \right)_{r_{ij}=0} r_{ij} + \mathcal{O}(r_{ij}^2). \quad (6)$$

Hence the antiparallel cusp condition is equivalent to the requirement that

$$\left( \frac{\partial J}{\partial r_{ij}} \right)_{r_{ij}=0} = \frac{1}{2}, \quad (7)$$

where  $r_{ij}$ ,  $r_i$ , and  $r_j$  are treated as independent variables.

#### B. The parallel-spin electron-electron cusp condition

Suppose now that the approaching electrons  $i$  and  $j$  have parallel spins. The cusp condition<sup>10</sup> is

$$\left( \frac{\partial \Psi_{1m}}{\partial r_{ij}} \right)_{r_{ij}=0} = \frac{1}{4} (\Psi_{1m})_{r_{ij}=0}, \quad (8)$$

where  $\Psi_{1m}$  is the  $r_{ij} Y_{1m}$  component of  $\Psi$  and  $Y_{lm}$  is the  $(l, m)$ th spherical harmonic.

Let us expand  $\Psi$  about  $\mathbf{r}_{ij}=0$ .  $D$  is an odd function of  $\mathbf{r}_{ij}$ ; hence we obtain

$$\begin{aligned} \Psi = \exp[J_{r_{ij}=0}] & \left\{ 1 + \left( \frac{\partial J}{\partial r_{ij}} \right)_{r_{ij}=0} r_{ij} \right. \\ & + \left[ \left( \frac{\partial J}{\partial r_i} \right) - \left( \frac{\partial J}{\partial r_j} \right) \right]_{r_{ij}=0} \delta r_i + \mathcal{O}(r_{ij}^2) \left. \right\} \\ & \times [(\nabla_{ij} D)_{r_{ij}=0} \cdot \mathbf{r}_{ij} + \mathcal{O}(r_{ij}^3)]. \end{aligned} \quad (9)$$

The change in the electron-nucleus distance when the electron separation  $r_{ij}$  is increased from zero is  $\delta r_i = r_{ij} \cos(\theta_i)/2 + \mathcal{O}(r_{ij}^2)$ , where  $\theta_i$  is the angle between  $\mathbf{r}_{ij}$  and  $\mathbf{r}_i$ . The  $r_{ij}Y_{1m}$  component of  $\Psi$  is therefore

$$\begin{aligned} \Psi_{1m} = \exp[J_{r_{ij}=0}] & [(\nabla_{ij} D)_{r_{ij}=0} \cdot \mathbf{r}_{ij}]_{1m} \\ & \times \left[ 1 + \left( \frac{\partial J}{\partial r_{ij}} \right)_{r_{ij}=0} r_{ij} + \mathcal{O}(r_{ij}^2) \right], \end{aligned} \quad (10)$$

where  $[X]_{1m}$  denotes the  $r_{ij}Y_{1m}$  component of  $X$ . So the parallel-spin cusp condition of Eq. (8) is equivalent to the requirement that

$$\left( \frac{\partial J}{\partial r_{ij}} \right)_{r_{ij}=0} = \frac{1}{4}, \quad (11)$$

where  $r_{ij}$ ,  $r_i$ , and  $r_j$  are treated as independent variables.

### C. The electron-nucleus cusp condition

Now consider the cusp condition that must be satisfied as electron  $i$  approaches a nucleus of atomic number  $Z$ . The coordinates of all other electrons are omitted. The spherical average of  $\Psi(\mathbf{r}_i)$  about the nucleus  $\bar{\Psi}(r_i)$  must obey<sup>10</sup>

$$\left( \frac{\partial \bar{\Psi}}{\partial r_i} \right)_{r_i=0} = -Z\bar{\Psi}_{r_i=0}. \quad (12)$$

By similar arguments to those given for the antiparallel electron-electron cusp condition, if the Slater determinant is continuously differentiable at the nucleus then the Jastrow factor must satisfy

$$\left( \frac{\partial J}{\partial r_i} \right)_{r_i=0} = -Z. \quad (13)$$

Note that if the Slater part of the wave function satisfies the electron-nucleus cusp condition, or if a nondivergent pseudopotential is used, then the Jastrow factor is required to be cuspless at the nuclei: it should satisfy Eq. (13) with  $Z=0$ .

## IV. THE BEHAVIOR OF THE LOCAL ENERGY AT COALESCENCE POINTS

### A. Continuity at antiparallel-spin coalescence points

The Slater-Jastrow wave function in the vicinity of an antiparallel-spin coalescence point can be written as

$$\Psi(\mathbf{r}_{ij}) = \exp[u(r_{ij})]S(\mathbf{r}_{ij}), \quad (14)$$

where  $S$  is the Slater wave function multiplied by the terms in the Jastrow factor that are analytic at the coalescence point

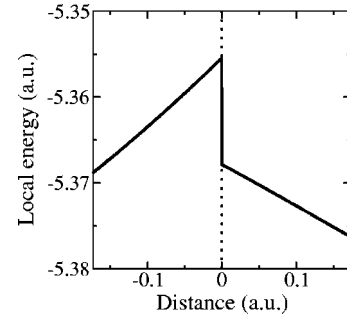


FIG. 1. Local energy plotted against the position of an electron as it is moved along a straight line through another electron of the opposite spin in  $\text{SiH}_4$ . The dotted line indicates the location of the stationary electron.

and we have assumed there are no  $f$  terms in the Jastrow factor. Assuming that  $u$  satisfies the Kato cusp condition of Eq. (7), the local energy can be shown to be

$$\begin{aligned} E_L(\mathbf{r}_{ij}) &= -\frac{\nabla_{ij}^2 \Psi}{\Psi} + \frac{1}{r_{ij}} + E_{L0} \\ &= -\frac{1}{4} - 3 \left( \frac{d^2 u}{dr_{ij}^2} \right)_{r_{ij}=0} - \frac{\nabla_{ij}^2 S}{S} + E_{L0} - \frac{(\nabla_{ij} S)_{r_{ij}=0} \cdot \mathbf{r}_{ij}}{S_{r_{ij}=0} \times r_{ij}} \\ &\quad + \mathcal{O}(r_{ij}), \end{aligned} \quad (15)$$

where the  $E_{L0}$  and  $-S^{-1}\nabla_{ij}^2 S$  terms are continuous at the coalescence point.

Satisfying the cusp condition removes the divergence in the local energy at the coalescence point, irrespective of the angle at which the electrons approach. However, the  $\mathcal{O}(r_{ij}^0)$  term in Eq. (15) *does* depend on the direction of approach. The local energy therefore has a point discontinuity at antiparallel-spin coalescence points. This behavior is illustrated in Fig. 1.

There is a similar discontinuity in the local energy at nuclei when the electron-nucleus cusp condition is enforced. If, on the other hand, the no-cusp condition is enforced at the center of a pseudoatom, there is no discontinuity in the local energy.

### B. Continuity at parallel-spin coalescence points

Now consider a parallel-spin coalescence point. Again the wave function may be written in the form of Eq. (14) in the vicinity of the coalescence point, but this time  $S$  is an odd function of  $\mathbf{r}_{ij}$ . If the Kato cusp condition of Eq. (11) is satisfied by  $u$ , the local energy is

$$E_L = -\frac{1}{16} - 5 \left( \frac{d^2 u}{dr_{ij}^2} \right)_{r_{ij}=0} - \frac{\nabla_{ij}^2 S}{S} + E_{L0} + \mathcal{O}(r_{ij}). \quad (16)$$

The  $-S^{-1}\nabla_{ij}^2 S$  term is discontinuous at a parallel-spin coalescence point, giving a point discontinuity in the local energy. In spite of this, the local energy is continuous when one electron is moved along a straight line through another of the same spin because of the symmetry of the local energy with respect to exchanges of parallel-spin electrons. This behavior is illustrated in Fig. 2.

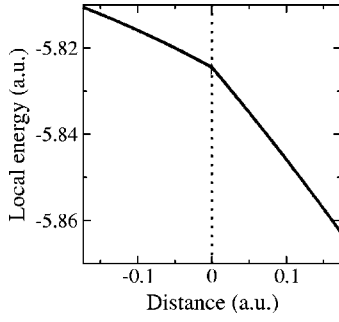


FIG. 2. Local energy plotted against the position of an electron as it is moved along a straight line through another electron of the same spin in  $\text{SiH}_4$ . The dotted line indicates the location of the stationary electron.

### C. Further coalescence conditions

Rassolov and Chipman<sup>11</sup> have demonstrated that, at the coalescence point of two antiparallel-spin electrons  $i$  and  $j$

$$\left( \frac{\partial^3 \hat{\Psi}}{\partial r_{ij}^3} \right)_{r_{ij}=0} = \left( \frac{\partial^2 \hat{\Psi}}{\partial r_{ij}^2} \right)_{r_{ij}=0} - \frac{\hat{\Psi}_{r_{ij}=0}}{8}, \quad (17)$$

while for two parallel-spin electrons

$$\left( \frac{\partial^3 \Psi_{1m}}{\partial r_{ij}^3} \right)_{r_{ij}=0} = \frac{7}{12} \left( \frac{\partial^2 \Psi_{1m}}{\partial r_{ij}^2} \right)_{r_{ij}=0} - \frac{(\Psi_{1m})_{r_{ij}=0}}{48}. \quad (18)$$

These cusp conditions are difficult to apply to Slater-Jastrow wave functions because Eqs. (17) and (18) involve the Slater determinant as well as the Jastrow factor. If we assume that  $(\partial^2 \hat{S} / \partial r_{ij}^2)_{r_{ij}=0} = 0$ , where  $\hat{S}$  is the spherical average of  $S$  about an antiparallel-spin coalescence point, then we can derive an approximate condition on the antiparallel-spin  $u$  term in the Jastrow factor.<sup>12</sup> Likewise, if we assume that  $(\partial^2 S_{1m} / \partial r_{ij}^2)_{r_{ij}=0} = 0$ , where  $S_{1m}$  is the  $r_{ij} Y_{1m}$  component of  $S$  about a parallel-spin coalescence point, then we can derive an approximate condition on the parallel-spin  $u$  term. Imposing these additional conditions was not found to be of any benefit in practice.

The Rassolov-Chipman conditions can be derived by demanding that  $\hat{\Psi}^{-1} \hat{H} \hat{\Psi}$  and  $(\Psi_{1m} r_{ij} Y_{1m})^{-1} \hat{H} (\Psi_{1m} r_{ij} Y_{1m})$  are cusplike at antiparallel- and parallel-spin coalescence points, respectively.<sup>12</sup> There seems little point in attempting to apply the Rassolov-Chipman conditions unless one has already ensured that the local energy itself is continuous at coalescence points. A continuous local energy can be achieved in a two-electron atom by using a trial wave function based upon the Fock expansion.<sup>13</sup> However, it is unlikely that a practical method for eliminating the local-energy discontinuities in larger systems will be found.

## V. NEW JASTROW FACTOR

### A. The $u$ , $\chi$ , and $f$ terms

For the  $u$  term we use an expression which is a variation on the form we have used for a number of years<sup>14</sup> and con-

sists of a complete power expansion in  $r_{ij}$  up to order  $r_{ij}^{C+N_u}$  which satisfies the Kato cusp conditions at  $r_{ij}=0$ , goes to zero at the cutoff length  $r_{ij}=L_u$ , and has  $C-1$  continuous derivatives at  $L_u$ :

$$u(r_{ij}) = (r_{ij} - L_u)^C \Theta(L_u - r_{ij}) \times \left( \alpha_0 + \left[ \frac{\Gamma_{ij}}{(-L_u)^C} + \frac{\alpha_0 C}{L_u} \right] r_{ij} + \sum_{l=2}^{N_u} \alpha_l r_{ij}^l \right), \quad (19)$$

where  $\Theta$  is the Heaviside function and  $\Gamma_{ij}=1/2$  if electrons  $i$  and  $j$  have opposite spins and  $\Gamma_{ij}=1/4$  if  $i$  and  $j$  have the same spin. In this expression  $C$  determines the behavior at the cutoff length. If  $C=2$ , the gradient of  $u$  is continuous but the second derivative and hence the local energy is discontinuous, and if  $C=3$  then both the gradient of  $u$  and the local energy are continuous.

The form of  $\chi$  is similarly related to our earlier work

$$\chi_I(r_{il}) = (r_{il} - L_{\chi I})^C \Theta(L_{\chi I} - r_{il}) \times \left( \beta_{0I} + \left[ \frac{-Z_I}{(-L_{\chi I})^C} + \frac{\beta_{0I} C}{L_{\chi I}} \right] r_{il} + \sum_{m=2}^{N_{\chi}} \beta_{mI} r_{il}^m \right). \quad (20)$$

It may be assumed that  $\beta_{mI} = \beta_{mJ}$  where  $I$  and  $J$  are equivalent ions. The term involving the ionic charge  $Z_I$  enforces the electron-nucleus cusp condition.

The expression for  $f$  is the most general expansion of a function of  $r_{ij}$ ,  $r_{il}$ , and  $r_{jl}$  that is cusplike at the coalescence point and goes smoothly to zero when either  $r_{il}$  or  $r_{jl}$  reach cutoff lengths

$$f_I(r_{il}, r_{jl}, r_{ij}) = (r_{il} - L_{fI})^C \Theta(L_{fI} - r_{il}) \times (r_{jl} - L_{fI})^C \Theta(L_{fI} - r_{jl}) \times \sum_{l=0}^{N_{fI}^{eN}} \sum_{m=0}^{N_{fI}^{eN}} \sum_{n=0}^{N_{fI}^{ee}} \gamma_{lmnI} r_{il}^l r_{jl}^m r_{ij}^n. \quad (21)$$

Various restrictions are placed on  $\gamma_{lmnI}$ . To ensure the Jastrow factor is symmetric under electron exchanges we demand that  $\gamma_{lmnI} = \gamma_{mlnI} \forall I, m, l, n$ . If ions  $I$  and  $J$  are equivalent then we demand that  $\gamma_{lmnI} = \gamma_{lmnJ}$ . The condition that the  $f$  term has no electron-electron cusps is

$$\left( \frac{\partial f}{\partial r_{ij}} \right)_{r_{ij}=0, r_{il}=r_{jl}} = 0, \quad (22)$$

which implies that

$$\sum_{l=0}^{N_{fI}^{eN}} \sum_{m=0}^{N_{fI}^{eN}} \gamma_{lm1I} r_{il}^{l+m} (r_{il} - L_{fI})^{2C} = 0 \quad (23)$$

for all  $r_{il}$ . Hence,  $\forall k \in \{0, \dots, 2N_{fI}^{eN}\}$ , we must have

$$\sum_{l,m:l+m=k} \gamma_{lm1I} = 0. \quad (24)$$

The condition that the  $f$  term has no electron-nucleus cusps is



$$\left(\frac{\partial f}{\partial r_{il}}\right)_{\substack{r_{il}=0 \\ r_{ij}=r_{jl}}} = 0, \quad (25)$$

which gives

$$\sum_{m=0}^{N_{fl}^{ee}} \sum_{n=0}^{N_{fl}^{ee}} (C\gamma_{0mnl} - L_{fl}\gamma_{1mnl}) \times (-L_{fl})^{C-1} r_{jl}^{m+n} (r_{jl} - L_{fl})^C = 0, \quad (26)$$

for all  $r_{jl}$ . We therefore require that,  $\forall k' \in \{0, \dots, N_{fl}^{ee} + N_{fl}^{ee}\}$

$$\sum_{m,n:m+n=k'} (C\gamma_{0mnl} - L_{fl}\gamma_{1mnl}) = 0. \quad (27)$$

The method by which we impose the various constraints is described in detail in Appendix A.

### B. The $p$ and $q$ terms

The  $p$  term takes the cusplless form

$$p(\mathbf{r}_{ij}) = \sum_A a_A \sum_{\mathbf{G}_A^+} \cos(\mathbf{G}_A \cdot \mathbf{r}_{ij}), \quad (28)$$

where the  $\{\mathbf{G}_A\}$  are the reciprocal lattice vectors of the simulation cell belonging to the  $A$ th star of vectors that are equivalent under the full symmetry group of the Bravais lattice, and “+” means that, if  $\mathbf{G}_A$  is included in the sum,  $-\mathbf{G}_A$  is excluded.

For systems with inversion symmetry the  $q$  term takes the cusplless form

$$q(\mathbf{r}_i) = \sum_B b_B \sum_{\mathbf{G}_B^+} \cos(\mathbf{G}_B \cdot \mathbf{r}_i), \quad (29)$$

where the  $\{\mathbf{G}_B\}$  are the reciprocal lattice vectors of the primitive unit cell belonging to the  $B$ th star of vectors that are equivalent under the space-group symmetry of the crystal, and the “+” means that, if  $\mathbf{G}_B$  is included in the sum,  $-\mathbf{G}_B$  is excluded. In this work  $q$  has only been used for systems with inversion symmetry. Plane-wave expansions of electron position can also be used for systems without inversion symmetry.<sup>15</sup>

### C. Cutting off terms in the Jastrow factor

To avoid unwanted derivative discontinuities in the wave function of a periodic system, the isotropic functions  $u$ ,  $\chi_I$ , and  $f_I$  must be cut off at a distance less than or equal to the Wigner-Seitz radius of the simulation cell. Furthermore, reasons of efficiency dictate that in particular  $f_I$  should be cut off at short distances in both finite and periodic systems. Suppose we wish to evaluate a Slater-Jastrow wave function for a number of systems of increasing size, where the number of electrons  $N$  is assumed to be proportional to the number of ions  $N_{\text{ions}}$ . If the cutoff lengths  $L_u$ ,  $L_\chi$ , and  $L_f$  are chosen to be proportional to the size of the system then the numbers of operations required to update  $u$  and  $\chi_I$  after each electron move are  $\mathcal{O}(N)$ . The cost of updating the  $f_I$  term is,

however, proportional to  $\mathcal{O}(N^2)$ , which is prohibitive in large systems. If we choose the cutoff lengths to be independent of the system size then each term can be updated in  $\mathcal{O}(1)$  operations.

The  $f_I$  term describes inhomogeneous correlations that are spherically symmetric about atom  $I$ . It does not seem likely that  $f_I$  could describe the inhomogeneity in correlations at points distant from atom  $I$  in systems with many atoms. Similarly we argue that the long-ranged part of the spherically symmetric  $\chi_I$  terms will not give useful variational freedom. For a fixed number of variable parameters we therefore expect that the best results will be obtained by cutting off the  $\chi_I$  and  $f_I$  terms at distances of roughly the size of atom  $I$ , so that the variational freedom in these terms is concentrated at short distances where it is most useful. The  $u$  term must describe both long- and short-ranged correlations and therefore we expect it to be long ranged. In our implementation we allow the cutoff lengths  $L_u$ ,  $L_\chi$ , and  $L_f$  to be varied, and these degrees of freedom are investigated in Secs. VIII–XI. In many cases the optimal value of  $L_u$  is approximately proportional to the system size and the optimal values of  $L_\chi$  and  $L_f$  are approximately independent of the system size, so that overall the cost of updating the Jastrow factor after each electron move scales as  $\mathcal{O}(N)$ .

As mentioned earlier, the value of  $C$  in Eqs. (19)–(21) determines the behavior of the Jastrow factor, and hence the local energy, at the cutoff lengths. Discontinuities in the local energy at the cutoff lengths may be harmful to optimization procedures, but the price paid for having a smoother local energy is a reduction in variational freedom.

## VI. FURTHER COMMENTS ON OUR JASTROW FACTOR

We have used power series in the interparticle distances rather than scaled variables, such as  $r_{ij}/(1+br_{ij})$ , which have been used by Boys and Handy<sup>16</sup> and others.<sup>17,18</sup> These scaled variables go to a constant at large  $r_{ij}$ , which is useful in finite systems. However, it is not clear whether scaled variables are helpful when the  $u$ ,  $\chi$ , and  $f$  terms are cut off at finite lengths, as they must be in periodic systems.

In our previous Jastrow factors<sup>14</sup> we used Chebyshev polynomials rather than the powers themselves. The ideas behind this were that (i) the Chebyshev series may be calculated to very high accuracy with double-precision arithmetic using recurrence relations and (ii) the optimal coefficients tend to be of a more uniform magnitude, which could be helpful within optimization procedures. However, we have found that the precision offered by a simple power series with double-precision arithmetic is perfectly acceptable up to an order of at least 20, and we have found no clear benefits from the use of Chebyshev polynomials within our current optimization procedures either. We have therefore chosen to use simple power series, which may be evaluated more rapidly than the corresponding Chebyshev series.

Our new Jastrow factor includes terms such as  $r_i h(r_j, r_{ij})$  and  $r_{ij} g(r_i, r_j)$ , where  $h$  and  $g$  are polynomials, which are absent in the Jastrow factors used by Schmidt and Moskowitz<sup>17,19</sup> and some other researchers. Such terms do not in general satisfy the cusp conditions on their own, but

certain linear combinations of them do and therefore they should be allowed to occur in the power series. The Jastrow factors used by Umrigar and co-workers<sup>18,20,21</sup> have included such terms for many years. We report tests of the importance of the terms neglected in the Schmidt-Moskowitz Jastrow factor in Sec. VIII.

As an option within our implementation we may try to reduce the extent to which  $f$  duplicates the  $u$  and  $\chi$  terms in the Jastrow factor by imposing the conditions  $\gamma_{00nl}=0$  for all  $n$  and  $\gamma_{l00l}=0$  for all  $l$ . Note, however, that the terms of  $f$  with  $l=m=0$  do not exactly correspond to  $u$ : they are electron-electron terms local to ion  $I$ . This variational freedom, which is investigated in Sec. VIII, may be used to describe correlations that occur on two different length scales, for example, in the core and valence electrons of an atom or the intra-atomic and interatomic electron correlations of weakly interacting atoms. The terms of  $f$  with  $m=n=0$  are less likely to give useful variational freedom. However, the use of duplication of  $u$  and  $\chi$  by  $f$  does not appear to cause any difficulties within our optimization procedure, even where  $L_\chi \simeq L_f$ .

The variable parameters appear linearly in our Jastrow factor, with the exception of the cutoff lengths  $L_u$ ,  $L_\chi$ , and  $L_{fj}$ . The least-squares function in an unweighted variance minimization is quartic in the linear parameters;<sup>8</sup> however, the dependence on the cutoff lengths is much more complicated, and they couple strongly to the other parameters. The use of linear parameters is found to be very advantageous in practice: up to ten times fewer Gauss-Newton iterations are required to converge to the minimum of the least-squares function when the cutoff lengths are fixed compared with when they are optimized.

Our Jastrow factor does not include logarithmic terms such as those motivated by the Fock expansion.<sup>22,23</sup> Although these terms have been used in highly accurate Hylleraas-expansion calculations for two-electron atoms<sup>13</sup> it should be noted that the most accurate calculations of this type performed to date have not included them.<sup>24</sup> It cannot therefore be necessary to include the logarithmic terms to obtain high accuracy.

Our Jastrow factor does not include terms involving three or more electrons. Of course, the repulsive Coulomb interaction and the antisymmetry of the wave function ensures that three or more electrons rarely come close to one another, so that such terms are expected to be small, and explicit tests by Huang *et al.*<sup>21</sup> support this view.

In our implementation it is possible to use different  $u$ ,  $f$ , and  $p$  functions for antiparallel, parallel spin-up, and parallel spin-down pairs of electrons, and different  $\chi$  and  $q$  functions for spin-up and spindown electrons. We investigate the effect of using the different possible spin dependences in the Jastrow factor for a partially polarized system in Sec. X. Note that if different  $u$  functions are used for parallel- and antiparallel-spin pairs of electrons then both of the Kato cusp conditions are satisfied, but if the same  $u$  function is used for all pairs of electrons then only the antiparallel-spin cusp condition is satisfied. The use of a Jastrow factor that is not symmetric with respect to exchanges of electrons of opposite spin generally produces a trial wave function that is not an eigenfunction of the spin operator  $\hat{S}^2$ , even though the

ground-state wave function must be an eigenfunction of  $\hat{S}^2$ . An investigation into this *spin-contamination* effect has been carried out by Huang *et al.*,<sup>25</sup> who found that highly optimized wave functions suffer from relatively little spin contamination.

## VII. SPECIFICATION OF THE JASTROW FACTOR AND THE MEASURE OF ACCURACY

In the tests reported here the parameter  $C$ , which determines the behavior at the cutoff lengths, takes the values  $C=2$  or  $C=3$ . The terms included in  $u$  are specified by  $N_u$ , those in  $\chi$  by  $N_\chi$ , and those in  $f$  by  $N_f^{eN}$  and  $N_f^{ee}$ . In each case  $N_\chi$  is the same for all species of atom present, and likewise for  $N_f^{eN}$  and  $N_f^{ee}$ . Spin dependences in  $u$ ,  $\chi$ ,  $f$ ,  $p$ , and  $q$  are specified by  $S_u$ ,  $S_\chi$ ,  $S_f$ ,  $S_p$ , and  $S_q$ , where  $S_u=0$  denotes that the same  $u$  function is used for parallel- and antiparallel-spin pairs,  $S_u=1$  denotes that different functions are used for parallel- and antiparallel-spin pairs, and  $S_u=2$  denotes that different functions are used for parallel spin-up, parallel spin-down, and antiparallel-spin pairs.  $S_\chi=0$  denotes that the same  $\chi$  function is used for spin-up and spin-down electrons while  $S_\chi=1$  if they are allowed to be different, etc. Duplication of the  $r_{ij}$  terms in  $u$  and the  $r_i$  terms in  $\chi$  by  $f$  is denoted by  $D=true$ . The terms included in the plane-wave expansions  $p$  and  $q$  are determined by the number of stars of  $\mathbf{G}$  vectors included,  $N_p$  and  $N_q$ . If  $SMJ=true$  then only the  $f$  terms contained in the Schmidt-Moskowitz Jastrow (SMJ) factor are used (i.e., the terms proportional to  $r_{ij}$  and  $r_i$  are omitted). In each case we will specify the relevant descriptors and give the total number of optimized parameters in the Jastrow factor  $N_T$ . The cutoff lengths are included in the count of parameters.

Unless otherwise stated, the  $\chi_I$  functions were chosen to be cusplless at the nuclei (i.e.,  $Z_I=0$  in each case) because nondivergent pseudopotentials were used or, where the full Coulomb potential was used, the orbitals satisfied the electron-nucleus cusp condition.

To initiate the optimization procedure one must select a set of configurations from a suitable probability distribution. We have found that the distribution obtained from the square of the Slater part of the wave function is normally an excellent starting point; indeed our results suggest that it may be preferable to the “self-consistent” approach of updating the distribution to include the latest estimate of the Jastrow factor. It should be noted that if one sets all of the variable parameters in our Jastrow factor to zero, the resulting wave function can be very poor, often giving energies which are higher than that obtained using the Slater part only.

We measure the accuracy of a Jastrow factor by the percentage of the DMC correlation energy retrieved within VMC, i.e.,

$$\eta = \frac{E_{\text{HF}} - E_{\text{VMC}}}{E_{\text{HF}} - E_{\text{DMC}}} \times 100\% , \quad (30)$$

where  $E_{\text{HF}}$  is the energy obtained with the Slater determinants only,  $E_{\text{VMC}}$  is the VMC energy obtained with the Slater-Jastrow wave function, and  $E_{\text{DMC}}$  is the DMC energy.

TABLE I. Optimized Jastrow factors and VMC energies for He. The HF energy is  $-2.86167999$  a.u., the exact energy is  $-2.903724$  a.u., and the DMC energy is within error bars of the exact value. In each case  $S_u=S_f=S_\chi=0$ ,  $N_u=N_\chi \equiv N_{u,\chi}$ , and  $N_f^{eN}=N_f^{ee} \equiv N_F^{eN,ee}$ .

$C$	$N_{u,\chi}$	$N_f^{eN,ee}$	$D$	SMJ	$N_T$	$E_{\text{VMC}}$ (a.u.)	$\eta$	$\sigma_E^2$ (a.u.)
3	8	0			18	$-2.900010(9)$	91.17(2)%	0.0237(4)
3	6	3	T	T	33	$-2.903555(2)$	99.598(5)%	0.002450(6)
2	8	3	F	F	40	$-2.903596(2)$	99.696(5)%	0.00246(6)
3	6	3	F	F	36	$-2.903660(3)$	99.848(7)%	0.00083(1)
3	6	3	T	F	41	$-2.903693(1)$	99.926(2)%	0.000653(4)

The DMC method gives the energy corresponding to a perfect Jastrow factor: see Sec. I. In this work the orbitals in the Slater determinants were kept fixed and we only optimized the Jastrow factor. Under these conditions  $\eta$  is an appropriate measure of the accuracy of Jastrow factors. We also report the variance of the local energy  $\sigma_E^2$  for each Jastrow factor tested. The energy variance is the quantity that determines the size of the statistical error bars for a given computational effort in QMC calculations. Furthermore, it is the object that we actually minimize when optimizing the Jastrow factor.

### VIII. EXAMPLE I: HE AND NE ATOMS

#### A. Two-electron atoms

Extremely accurate energies are available for the two-electron He and  $\text{Ne}^{8+}$  atoms from variational calculations using Hylleraas expansions<sup>24</sup> and other methods. It is straightforward to show that the exact ground-state wave function of a two-electron atom is a nodeless function of  $r_1$ ,  $r_2$ , and  $r_{12}$ . It should therefore be possible to obtain very accurate results by including  $f(r_1, r_2, r_{12})$  terms in the Jastrow factor. As the ground-state wave function is nodeless, the DMC energy should equal the exact (nonrelativistic and infinite-nuclear-mass) energy, apart from statistical errors and biases due to the use of finite time steps and populations. We used orbitals derived from numerical integrations of the Hartree-Fock (HF) equations on fine radial grids.

Tables I and II show variational energies of optimized Jastrow factors for He and  $\text{Ne}^{8+}$ . When using  $u$  and  $\chi$  functions only we obtain 91.17(2)% (He) and 93.64(11)% ( $\text{Ne}^{8+}$ )

TABLE II. Optimized Jastrow factors and VMC energies for  $\text{Ne}^{8+}$ . The HF energy is  $-93.86111347$  a.u., the exact energy is  $-93.906806$  a.u., and the DMC energy is within error bars of the exact value. Duplication of  $u$  and  $\chi$  by  $f$  is permitted. In each case  $S_u=S_f=S_\chi=0$ ,  $N_u=N_\chi \equiv N_{u,\chi}$ , and  $N_f^{eN}=N_f^{ee} \equiv N_f^{eN,ee}$ .

$C$	$N_{u,\chi}$	$N_f^{eN,ee}$	SMJ	$N_T$	$E_{\text{VMC}}$ (a.u.)	$\eta$	$\sigma_E^2$ (a.u.)
2	4	0		10	$-93.90387(3)$	93.57(7)%	0.645(4)
2	8	0		18	$-93.90390(5)$	93.6(1)%	0.645(4)
3	8	0		18	$-93.90390(5)$	93.6(1)%	0.645(1)
2	4	3	T	29	$-93.90672(1)$	99.81(2)%	0.0810(3)
2	4	3	F	37	$-93.90672(2)$	99.81(4)%	0.0138(8)
3	6	3	F	41	$-93.906801(6)$	99.99(1)%	0.00276(7)

of the correlation energy, but when we add an  $f$  term we obtain nearly 100% of the correlation energy. Elimination of the terms of the forms  $r_i h(r_j, r_{ij})$  and  $r_{ij} g(r_i, r_j)$  leads to an expression containing the same powers as the Jastrow factor of Schmidt and Moskowitz.<sup>17,19</sup> The additional terms are unimportant in He and  $\text{Ne}^{8+}$ . The results are not strongly dependent on whether  $C=2$  or 3, or whether duplication of the terms in  $u$  and  $\chi$  by those in  $f$  is prevented or not.

Our results for  $\text{Ne}^{8+}$  are better than our results for He, both with and without the  $(r_i, r_j, r_{ij})$  terms in the Jastrow factor. This is to be expected, because the electron-electron interaction is a smaller perturbation in  $\text{Ne}^{8+}$ , and hence correlation effects are less significant.

Using a Jastrow factor consisting of a fourth-order Padé function of scaled variables, Umrigar *et al.*<sup>7</sup> obtained a VMC energy of  $-2.903726(4)$  a.u. for He, so they were able to retrieve 100% of the correlation energy in this case. Making use of scaled variables, instead of cutting off the Jastrow factor at a finite range, would therefore appear to be beneficial in the special case of two-electron atoms.

In general we find that the Jastrow factors which recover a large fraction of the correlation energy have a correspondingly low variance. However, the variance obtained for  $\text{Ne}^{8+}$  using SMJ=true is surprisingly high, even though the variational energy is about the same as the corresponding result in which the full variational freedom of  $f$  is used.

#### B. All-electron Ne and pseudo-Ne

The results of optimizing different Jastrow factors for the all-electron Ne atom are given in Table III. The importance of the  $f$  terms is clear: less than 60% of the correlation energy can be retrieved using only  $u$  and  $\chi$ , whereas more than 90% can be retrieved if  $f$  terms are used as well.

We find that using  $C=2$  gives slightly better results than  $C=3$ : it does not cause our optimization procedure any difficulties, and the extra variational freedom can be exploited in this case. The discontinuities in the local energy do not appear to cause any population-control problems for the DMC algorithm either.

The optimal values of the cutoff lengths  $L_u$ ,  $L_\chi$ , and  $L_f$  lie between 2 and 3 a.u. in most cases. In our best wave functions  $L_u$  is the longest of the three. Where  $\chi$  is absent, however,  $f$  has the greatest cutoff length. We tried optimizing more than one  $f$  function in order to allow separate  $(r_i, r_j, r_{ij})$  correlations for the core and valence electrons, but this did not lower the variational energy.



TABLE III. Optimized Jastrow factors and VMC energies for all-electron Ne. The HF energy is  $-128.54709807$  a.u., the exact energy is  $-128.9376$  a.u. (Refs. 26 and 27) and our DMC energy is  $-128.9238(7)$  a.u. In all cases  $S_u=1$ ,  $S_\chi=0$ ,  $N_u=8$ , and  $N_f^{eN}=N_f^{ee} \equiv N_f^{eN,ee}$ .

$C$	$N_\chi$	$N_f^{eN,ee}$	$S_f$	$D$	SMJ	$N_T$	$E_{\text{VMC}}$ (a.u.)	$\eta$	$\sigma_E^2$ (a.u.)
2	8	0				26	$-128.757(9)$	56(2)%	3.17(6)
2	0	2	0	F	F	23	$-128.781(9)$	62(2)%	3.2(1)
3	0	2	0	T	F	26	$-128.850(7)$	80(2)%	2.14(6)
2	0	2	0	T	F	26	$-128.863(7)$	84(2)%	2.2(1)
2	0	3	0	F	F	39	$-128.868(7)$	85(2)%	2.03(1)
2	4	3	0	T	T	41	$-128.876(2)$	87.3(6)%	1.92(1)
2	0	3	0	T	F	44	$-128.877(6)$	88(2)%	1.49(3)
2	4	3	0	T	F	49	$-128.886(2)$	90.0(6)%	1.27(2)
2	4	3	1	T	F	75	$-128.8983(2)$	93.2(2)%	1.12(2)

If  $N_f^{eN}=N_f^{ee}=2$  and  $\chi$  is absent then it is important to allow  $f$  to duplicate  $u$  and  $\chi$ . 62(2)% of the correlation energy is retrieved when duplication is disallowed whereas 84(2)% is retrieved when duplication is permitted. However, the difference is far less pronounced when  $N_f^{eN}=N_f^{ee}=3$ : about 85% is retrieved irrespective of whether duplication is allowed. Using  $N_f^{eN}=N_f^{ee}=2$  and allowing duplication of  $u$  and  $\chi$  gives a more efficient parametrization of the Jastrow factor, for the number of parameters is substantially less than is the case when  $N_f^{eN}=N_f^{ee}=3$ . In these calculations the optimal cutoff length of  $f$  ( $L_f \approx 3.7$  a.u.) is greater than that of  $u$  ( $L_u \approx 1.0$  a.u.). Isolated atoms are a special case in which the  $\chi$  function can be long ranged. In the absence of  $\chi$ ,  $f$  is forced to be long ranged so that it can describe the electron-nucleus correlations. Hence  $u$ , rather than  $f$ , has to describe all the short-ranged electron-electron correlations.

We obtain slightly better results when we include the terms in  $f$  that are neglected in the Schmidt-Moskowitz Jastrow factor. The VMC energy is fairly insensitive to the spin dependence of  $f$ .

We have investigated whether it is better to include the electron-nucleus cusp in the Jastrow factor or in the orbitals in the Slater wave function. Calculations were carried out using orbitals expanded in a Gaussian basis set, generated by the CRYSTAL code.<sup>28</sup> The  $\chi$  term in the Jastrow factor satisfied the electron-nucleus cusp condition. The results obtained were significantly poorer than those shown in Table III. In order to get reasonable variational energies, a very large number of  $\chi$  parameters was required, with  $N_\chi \geq 15$ . Even with  $N_\chi=15$ , only about 25% of the correlation energy was retrieved. It is clearly preferable to use orbitals that satisfy the electron-nucleus cusp condition.

There is a significant fixed-node error in the DMC energy: our DMC energy is 0.0138(7) a.u. higher than the exact non-relativistic ground-state energy.<sup>26,27</sup> We have verified that population-control biases are negligibly small and we have performed an extrapolation to zero time step, so the only remaining bias in our DMC energy is the fixed-node error. The best all-electron VMC energy reported in the literature is that of Huang *et al.*,<sup>21</sup> who optimized parameters in their orbitals at the same time as their Jastrow factor, giving them extra variational freedom, including the opportunity to re-

duce the fixed-node error. Using a Jastrow factor containing the same types of correlation as ours (electron-electron, electron-nucleus, and electron-electron-nucleus), and optimizing the orbitals as well as the Jastrow factor, Huang *et al.* obtain a VMC energy of  $-128.9008(1)$  a.u., which is only slightly lower than our best energy of  $-128.8983(2)$  a.u.

The results of optimizing Jastrow factors for pseudo-Ne are shown in Table IV.  $f$  is much less important in the pseudo-atom than in all-electron Ne. 86.3(5)% of the correlation energy is retrieved using  $u$  and  $\chi$  only, while 95.7(4)% is retrieved when  $f$  is used too. A greater fraction of the correlation energy can be retrieved in the pseudoatom than in the all-electron atom.

### IX. EXAMPLE II: SiH<sub>4</sub> MOLECULE

We used a bond length of 2.8289 a.u. for the SiH<sub>4</sub> (silane) molecule, in which the Si<sup>4+</sup> ion was represented by a relativistic HF pseudopotential<sup>30</sup> and the full Coulomb potential was used for the hydrogen nuclei. The orbitals forming the Slater determinants were obtained from HF calculations using a large Gaussian basis set and the GAUSSIAN code.<sup>31</sup>

Results for some of the Jastrow factors tested for SiH<sub>4</sub> are given in Table V. We find that a large fraction of the correlation energy can be obtained using rather simple Jastrow factors. Using  $u$  and  $\chi$  functions only, and with a total of only seven parameters, we are able to obtain almost 90% of the correlation energy. Our best Jastrow factors obtain about

TABLE IV. Optimized Jastrow factors and VMC energies for pseudo-Ne. An HF pseudopotential was used to represent the Ne<sup>8+</sup> ion (Ref. 29). The HF energy is  $-34.6105$  a.u. and the DMC energy is  $-34.9220(4)$  a.u. In all cases  $C=2$ ,  $S_u=1$ ,  $S_\chi=S_f=0$ ,  $N_u=N_\chi=8$ , and  $N_f^{eN}=N_f^{ee} \equiv N_f^{eN,ee}$ .

$N_f^{eN,ee}$	$D$	$N_T$	$E_{\text{VMC}}$ (a.u.)	$\eta$	$\sigma_E^2$ (a.u.)
0		26	$-34.879(1)$	86.2(3)%	0.74(2)
3	F	48	$-34.904(1)$	94.2(3)%	0.45(1)
2	F	32	$-34.905(1)$	94.5(3)%	0.51(3)
2	T	35	$-34.908(1)$	95.5(3)%	0.460(4)



TABLE V. Optimized Jastrow factors and VMC energies for SiH<sub>4</sub>. The HF energy is  $-6.118$  a.u. and the DMC energy is  $-6.3064(2)$  a.u. In each case  $S_u=1$ ,  $S_\chi=S_f=0$ ,  $N_u=N_\chi\equiv N_{u,\chi}$ , and  $N_f^e=N_f^e\equiv N_f^{eN,ee}$ . Duplication of  $u$  and  $\chi$  by  $f$  is prohibited.

$C$	$N_{u,\chi}$	$N_f^{eN,ee}$	$N_T$	$E_{\text{VMC}}$ (a.u.)	$\eta$	$\sigma_E^2$ (a.u.)
2	1	0	7	-6.284(2)	88(1)%	0.096(2)
2	12	0	51	-6.291(2)	92(1)%	0.066(4)
3	4	0	19	-6.291(2)	92(1)%	0.07(1)
3	4	2	31	-6.292(2)	92(1)%	0.08(2)
2	4	0	19	-6.293(2)	93(1)%	0.067(5)

93(1)% of the correlation energy. We find the optimal cutoff lengths ( $L_u \approx 10$  a.u. and  $L_{\chi\text{Si}} \approx L_{\chi\text{H}} \approx 5$  a.u.) to be fairly independent of  $N_u$  and  $N_\chi$ . There is no detectable benefit from going beyond  $N_u=N_\chi=4$ , or from introducing  $f$  functions. Both the results obtained and the behavior of the optimization procedure are very similar for  $C=2$  and  $C=3$ , so that in this case there is no benefit from having a continuous local

### X. EXAMPLE III: Ni ATOM AND NiO DIMER

We investigated the Ni atom and the NiO dimer with a bond length of 3.075 a.u., using HF pseudopotentials<sup>29</sup> to represent the Ni<sup>10+</sup> and O<sup>6+</sup> ions. The orbitals were obtained from HF calculations using a large Gaussian basis set and the CRYSTAL code.<sup>28</sup> We find that the  $f$  functions are significant for both the Ni atom (Table VI) and the NiO dimer (Table VII) in spite of the use of pseudopotentials.

Note that Ni and NiO are partially spin polarized, so that it may be advantageous to have different  $\chi$  functions for spin-up and spin-down electrons, and different  $u$  and  $f$  functions for parallel spin-up and parallel spin-down pairs of electrons, unlike the other systems studied in this work. Our results show that the spin dependences of the  $u$ ,  $\chi$ , and  $f$  functions do indeed have a significant effect on the quality of the wave functions for Ni and NiO, although including  $(r_i, r_j, r_{ij})$  terms in the Jastrow factor has a greater effect. An

TABLE VI. Optimized Jastrow factors and VMC energies for pseudo-Ni. The HF energy is  $-38.6670$  a.u. and the DMC energy is  $-39.2310(5)$  a.u. In each case  $C=3$ ,  $N_u=N_\chi=8$  and  $N_f^e=N_f^e\equiv N_f^{eN,ee}$ .

$N_f^{eN,ee}$	$S_u$	$S_\chi$	$S_f$	$D$	$N_T$	$E_{\text{VMC}}$ (a.u.)	$\eta$	$\sigma_E^2$ (a.u.)
0	1	0			26	-39.0598(4)	69.65(9)%	1.148(2)
0	1	1			34	-39.0673(4)	70.098(9)%	1.114(3)
0	2	1			42	-39.0719(4)	71.8(1)%	1.098(4)
3	1	0	0	F	48	-39.1045(3)	77.57(9)%	0.721(3)
3	1	1	0	F	56	-39.1074(3)	78.09(9)%	0.695(1)
3	2	1	0	F	64	-39.1231(3)	80.87(9)%	0.734(1)
3	2	1	1	F	85	-39.1247(3)	81.15(9)%	0.743(2)
3	2	1	2	F	106	-39.1247(3)	81.15(9)%	0.727(4)
3	2	1	2	T	121	-39.1469(2)	85.09(8)%	0.571(1)

TABLE VII. Optimized Jastrow factors and VMC energies for the NiO dimer. The HF energy is  $-54.31362$  a.u. and the DMC energy is  $-55.1558(6)$  a.u. In each case  $C=3$ ,  $S_u=1$ ,  $S_\chi=S_f\equiv S_{\chi,f}$ ,  $N_u=N_\chi=8$ , and  $N_f^e=N_f^e\equiv N_f^{eN,ee}$ . Duplication of  $u$  and  $\chi$  by  $f$  is forbidden.

$N_f^{eN,ee}$	$S_u$	$S_{\chi,f}$	$N_T$	$E_{\text{VMC}}$ (a.u.)	$\eta$	$\sigma_E^2$ (a.u.)
0	1	0	35	-54.9031(5)	69.99(8)%	2.144(4)
0	1	1	51	-54.9069(5)	70.45(8)%	2.132(4)
3	1	0	77	-54.9984(4)	81.31(7)%	1.521(3)
3	1	1	157	-55.0104(4)	82.74(8)%	1.451(2)
3	2	1	165	-55.0105(4)	82.75(8)%	1.439(2)

additional 1–2 % of the correlation energy can be retrieved when  $\chi$  and  $u$  are allowed to differ for spin-up and spin-down electrons. Using different  $f$  functions for antiparallel, parallel spin-up, and parallel spin-down pairs also lowers the variational energy slightly, although it greatly increases the number of parameters which have to be optimized. These calculations are the only ones for which we have retrieved less than 90% of the correlation energy.

### XI. EXAMPLE IV: SI SOLID

#### A. 16-atom simulation cell

We modeled crystalline Si in the diamond structure using a 16-atom, face-centered cubic simulation cell subject to periodic boundary conditions. We used a cubic lattice constant of 5.12966 a.u., and the Si<sup>4+</sup> ions were represented by pseudopotentials.<sup>30</sup> The orbitals were obtained from HF calculations using a large Gaussian basis set and the CRYSTAL code.<sup>28</sup> The results of optimizing the our Jastrow factor are shown in Table VIII.

$L_\chi$  adjusts itself to sizes of the order of the interatomic spacing ( $L_\chi \approx 6.4$  a.u., whereas the nearest-neighbor distance is 4.4424 a.u.), while  $L_u$  tends to the largest possible value, which is the Wigner-Seitz radius of the simulation cell (7.2544 a.u.). It is much easier to optimize the cutoff lengths when  $C=3$  than when  $C=2$ . It seems that the discontinuities in the local energy that are present when  $C=2$  cause serious

TABLE VIII. Optimized Jastrow factors and VMC energies for pseudo-Si (16-atom simulation cell). The HF energy is  $-7.63946$  a.u. per primitive cell and the DMC energy is  $-7.90600(6)$  a.u. per primitive cell. In each case  $S_u=S_p=1$ ,  $S_\chi=S_f=S_q=0$ ,  $N_u=N_\chi\equiv N_{u,\chi}$ ,  $N_f^e=N_f^e\equiv N_f^{eN,ee}$ , and  $N_p=N_q\equiv N_{p,q}$ . Duplication of  $u$  and  $\chi$  by  $f$  is allowed.

$C$	$N_{u,\chi}$	$N_f^{eN,ee}$	$N_{p,q}$	$N_T$	$E_{\text{VMC}}$ (a.u.)	$\eta$	$\sigma_E^2$ (a.u.)
2	1	0	0	5	-7.8583(5)	82.1(2)%	1.551(3)
2	10	0	0	32	-7.8714(4)	87.0(2)%	0.863(2)
3	1	0	0	5	-7.8761(4)	88.8(2)%	1.010(6)
3	10	0	0	32	-7.8809(4)	90.6(2)%	0.842(6)
3	4	0	5	29	-7.8816(3)	90.8(1)%	0.81(1)
3	4	2	0	23	-7.8832(3)	91.4(1)%	0.846(2)

problems for our optimization procedures. For example, a flexible Jastrow factor with  $C=2$  gives a lower variance, but higher energy, than a simple Jastrow factor with  $C=3$ . Such problems were not apparent in our calculations for atoms and small molecules. The discontinuities do not appear to lead to any problems within DMC calculations, however.

The use of the  $p$  and  $q$  terms does not bring about a statistically significant lowering of the VMC energy. The optimal value of  $L_u$  is a little less than the Wigner-Seitz radius when  $p$  terms are included, so the  $p$  term must describe the long-ranged correlations, as expected. The sinusoidal  $p$  and  $q$  functions are considerably more expensive to evaluate than the polynomial  $u$  and  $\chi$  functions, so it is anticipated that  $p$  and  $q$  will rarely be used in practice, except in strongly anisotropic systems such as graphite, where Prendergast *et al.*<sup>32</sup> have demonstrated that plane-wave expansions in the Jastrow factor have an important role to play.

### B. 54-atom simulation cell

Similar calculations to those reported in Sec. XI A were carried out using a 54-atom simulation cell. The HF energy is  $-7.6792$  a.u. per primitive cell and the DMC energy is  $-7.9555(2)$  a.u. per primitive cell. The VMC energy obtained using our Jastrow factor with  $N_u=N_\chi=4$ ,  $S_u=1$ , and  $S_\chi=0$  (giving 14 free parameters) is  $-7.9331(6)$  a.u. per primitive cell, so 91.9(2)% of the correlation energy is retrieved. A very similar fraction of the correlation energy is retrieved in the 16-atom cell (see Table VIII). The VMC energy variance is 2.92(5) a.u., so the variance per electron is about the same as for the 16-atom simulation cell. The optimal value of  $L_u$  is again equal to the Wigner-Seitz radius of the simulation cell (10.882 a.u.). The optimal value of  $L_\chi$  remains of order the atomic size, at  $L_\chi \approx 4.7$  a.u. These results indicate that cutting off the  $\chi$  function at sizes of order the inter-atomic spacing is a valuable improvement to the Jastrow factor.

## XII. CONCLUSIONS

We have developed and tested a form of Jastrow factor consisting of electron-electron ( $u$ ), electron-nucleus ( $\chi_I$ ), and electron-electron-nucleus ( $f_I$ ) terms, and additional electron-position-dependent ( $q$ ) and electron-electron-separation-dependent ( $p$ ) terms. The  $u$ ,  $\chi_I$ , and  $f_I$  terms are expanded in polynomials and are forced to go to zero at some cutoff radii. The  $p$  and  $q$  terms are expanded in plane waves. We have tested our Jastrow factor on atoms, molecules, and solids, including both all-electron and pseudopotential atoms. In most cases our VMC calculations have retrieved over 90% of the fixed-node correlation energy.

The variable parameters appear linearly in our Jastrow factor, except for the cutoff radii of  $u$ ,  $\chi_I$ , and  $f_I$ . The linearity in the variable parameters aids the computational efficiency of the optimization algorithm. We have found that it is often beneficial to make the local energy continuous at the cutoffs when optimizing the cutoff radii, but that a lower variational energy can be achieved in some cases when discontinuities are allowed at the cutoffs.

We have investigated the importance of terms in the Jastrow factor that were neglected by Schmidt and Moskowitz,<sup>17,19</sup> but we found them to be unimportant in the systems studied here. The electron-electron-nucleus  $f_I$  terms were found to be important in all-electron simulations of the He, Ne<sup>8+</sup>, and Ne atoms. The  $f_I$  terms are generally less important for pseudo-atoms than all-electron atoms. For example, they account for nearly 40% of the correlation energy in all-electron Ne, but only about 10% in pseudo-Ne. We found the  $f_I$  terms to be significant for pseudo-Ni and the NiO dimer, but they had little effect in SiH<sub>4</sub> or crystalline Si. The plane-wave terms  $p$  and  $q$  were found to be unimportant in crystalline Si.

We have found that it is preferable to use orbitals which satisfy the electron-nucleus cusp condition and to require the Jastrow factor to be cusplless at nuclei rather than to enforce the electron-nucleus cusp condition through the Jastrow factor. Overall we found that optimizing the cutoff radii is very important. In crystalline Si the cutoff for  $u$  adjusted itself to the largest possible value, which is the Wigner-Seitz radius of the simulation cell. This indicates that there are significant correlations in Si extending over many atoms. We have argued that the long-ranged parts of the  $\chi_I$  and  $f_I$  terms do not give useful variational freedom, and in support of this we found that the corresponding cutoffs adjusted themselves to sizes of the order of the interatomic spacing.

Although we have optimized our Jastrow factors by minimizing the variance of the energy, we have measured the accuracy of the Jastrow factors in terms of the variational energy itself. If we were to minimize the variational energy directly then we might get even better results. Recently there has been much interest in developing energy-minimization methods for optimizing trial wave functions,<sup>32-34</sup> and we intend to pursue this avenue further.

## ACKNOWLEDGMENTS

We thank Cyrus Umrigar for helpful discussions and John Trail for providing the all-electron HF orbitals and the pseudopotentials used in this work. Financial support was provided by the Engineering and Physical Sciences Research Council (EPSRC), UK. M.D.T. also thanks the Royal Society for providing financial support. Computational facilities were provided by the High Performance Computing Facility at the University of Cambridge.

## APPENDIX: IMPOSING THE CONSTRAINTS ON $\gamma$

Consider a particular set of ions  $I$  and a particular spin-pair type. To impose the symmetry of  $f$  under interchange of electrons we work with  $\gamma_{lmnl}$ , where  $l \geq m$ , and then complete the  $\gamma$  array by setting  $\gamma_{lmnl} = \gamma_{mlnl}$  for  $l < m$ . Let  $\mathbf{x}$  be a vector whose components are each of the  $\gamma$  coefficients with  $l \geq m$ . The remaining constraints (no-cusp conditions and, optionally, no-duplication-of- $u$ -and- $\chi$  conditions) may then be written in matrix form as  $A\mathbf{x}=0$ . The total number of constraints determines the number of rows of  $A$ .

There are  $2N_{fI}^e + 1$  constraints (one for each value of  $k = l+m$ ) associated with the imposition of the electron-

electron no-cusp condition. Using the symmetry of  $\gamma$ , Eq. (24) can be rewritten so that only elements of  $\mathbf{x}$  are involved:  $\forall k \in \{0, \dots, 2N_{fI}^{eN}\}$ ,

$$\sum_{\substack{l+m=k \\ l,m: l>m}} 2\gamma_{lm1I} + \sum_{l:2l=k} \gamma_{ll1I} = 0. \quad (\text{A1})$$

For each  $k$ , this equation defines a row of  $A$ .

There are  $N_{fI}^{eN} + N_{fI}^{ee} + 1$  constraints (one for each value of  $k' = l+n$ ) associated with the imposition of the electron-nucleus no-cusp condition. Equation (27) can be rewritten so that only elements of  $\mathbf{x}$  are involved:  $\forall k' \in \{0, \dots, N_{fI}^{eN} + N_{fI}^{ee}\}$ ,

$$C\gamma_{00k'I} - L_{fI}\gamma_{10k'I} + \sum_{\substack{l,n: l+n=k' \\ l \geq 1}} (C\gamma_{l0nI} - L_{fI}\gamma_{l1nI}) = 0. \quad (\text{A2})$$

For each  $k'$ , this equation defines a row of  $A$ .

If desired, there are  $N_{fI}^{ee}$  constraints imposed to prevent duplication of  $u$ . ( $\gamma_{00nI} = 0 \forall n$ .)

If desired, there are  $N_{fI}^{eN}$  constraints imposed to prevent duplication of  $\chi$ . ( $\gamma_{l00I} = 0 \forall l$ .)

Imposing the constraints reduces the number of independent variable parameters by rank ( $A$ ). The matrix  $A$  is transformed into its row-reduced echelon form by Gaussian elimination, which allows us to identify a suitable set of independent variable parameters.

- 
- <sup>1</sup>D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).  
<sup>2</sup>W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).  
<sup>3</sup>J. B. Anderson, J. Chem. Phys. **65**, 4121 (1976).  
<sup>4</sup>M. M. Hurley and P. A. Christiansen, J. Chem. Phys. **86**, 1069 (1987).  
<sup>5</sup>B. L. Hammond, P. J. Reynolds, and W. A. Lester, Jr., J. Chem. Phys. **87**, 1130 (1987).  
<sup>6</sup>L. Mitas, E. L. Shirley, and D. M. Ceperley, J. Chem. Phys. **95**, 3467 (1991).  
<sup>7</sup>C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988).  
<sup>8</sup>P. R. C. Kent, R. J. Needs, and G. Rajagopal, Phys. Rev. B **59**, 12 344 (1999).  
<sup>9</sup>R. J. Needs, M. D. Towler, N. D. Drummond, P. R. C. Kent, and A. J. Williamson, CASINO version 1.6 User Manual (University of Cambridge, Cambridge, 2002).  
<sup>10</sup>T. Kato, Commun. Pure Appl. Math. **10**, 151 (1957); R. T. Pack and W. B. Brown, J. Chem. Phys. **45**, 556 (1966).  
<sup>11</sup>V. A. Rassolov and D. M. Chipman, J. Chem. Phys. **104**, 9908 (1996).  
<sup>12</sup>N. D. Drummond, Ph.D. thesis, University of Cambridge, Cambridge, 2004.  
<sup>13</sup>C. R. Myers, C. J. Umrigar, J. P. Sethna, and J. D. Morgan III, Phys. Rev. A **44**, 5537 (1991).  
<sup>14</sup>A. J. Williamson, S. D. Kenny, G. Rajagopal, A. J. James, R. J. Needs, L. M. Fraser, W. M. C. Foulkes, and P. Maccallum, Phys. Rev. B **53**, 9640 (1996).  
<sup>15</sup>A. Malatesta, S. Fahy, and G. B. Bachelet, Phys. Rev. B **56**, 12 201 (1997).  
<sup>16</sup>S. F. Boys and N. C. Handy, Proc. R. Soc. London, Ser. A **310**, 63 (1969).  
<sup>17</sup>K. E. Schmidt and J. W. Moskowitz, J. Chem. Phys. **93**, 4172 (1990).  
<sup>18</sup>C. Filippi and C. J. Umrigar, J. Chem. Phys. **105**, 213 (1996).  
<sup>19</sup>J. W. Moskowitz and K. E. Schmidt, J. Chem. Phys. **97**, 3382 (1992).  
<sup>20</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, Berlin, 1988).  
<sup>21</sup>C.-J. Huang, C. J. Umrigar, and M. P. Nightingale, J. Chem. Phys. **107**, 3007 (1997).  
<sup>22</sup>V. A. Fock, D. Kngl. Norske Videnskab. Selsk. Forh. **31**, 138 (1958); **31**, 145 (1958).  
<sup>23</sup>J. D. Morgan III, Thermochim. Acta **69**, 181 (1986).  
<sup>24</sup>G. W. F. Drake, M. M. Cassar, and R. Nistor, Phys. Rev. A **65**, 054501 (2002).  
<sup>25</sup>C.-J. Huang, C. Filippi, and C. J. Umrigar, J. Chem. Phys. **108**, 8838 (1998).  
<sup>26</sup>E. R. Davidson, S. A. Hagstrom, S. J. Chakravorty, V. M. Umar, and C. F. Fischer, Phys. Rev. A **44**, 7071 (1991).  
<sup>27</sup>S. J. Chakravorty, S. R. Gwaltney, E. R. Davidson, F. A. Parpia, and C. F. Fischer, Phys. Rev. A **47**, 3649 (1993).  
<sup>28</sup>V. R. Saunders, R. Dovesi, C. Roetti, M. Causà, N. M. Harrison, R. Orlando, and C. M. Zicovich-Wilson, *CRYSTAL98 User's Manual* (University of Torino, Torino, 1998).  
<sup>29</sup>J. R. Trail and R. J. Needs (unpublished).  
<sup>30</sup>Y. Lee, Ph.D. thesis, University of Cambridge, Cambridge, 2002.  
<sup>31</sup>M. J. Frisch *et al.*, GAUSSIAN98, revision A.9 (Gaussian, Pittsburgh, PA, 1998).  
<sup>32</sup>D. Prendergast, D. Bevan, and S. Fahy, Phys. Rev. B **66**, 155104 (2002).  
<sup>33</sup>C. Filippi and S. Fahy, J. Chem. Phys. **112**, 3523 (2000).  
<sup>34</sup>X. Li, H. Zhang, and A. M. Rappe, J. Chem. Phys. **112**, 2560 (1999).