

Optimized wave functions for quantum Monte Carlo studies of atoms and solids

A. J. Williamson,* S. D. Kenny, and G. Rajagopal

Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, England

A. J. James

The Blackett Laboratory, Imperial College, Prince Consort Road, London SW7 2BZ, United Kingdom

R. J. Needs

Cavendish Laboratory, Madingley Road, Cambridge CB3 0HE, England

L. M. Fraser and W. M. C. Foulkes

The Blackett Laboratory, Imperial College, Prince Consort Road, London SW7 2BZ, United Kingdom

P. Maccallum

Edinburgh Parallel Computing Centre, James Clerk Maxwell Building, The King's Buildings, University of Edinburgh, Mayfield Road, Edinburgh EH9 3JZ, Scotland

(Received 18 August 1995)

Wave functions for the homogeneous electron gas, a germanium pseudosolid, and a germanium pseudoatom are optimized using the method of variance minimization. Forms for the Jastrow factor which are convenient to optimize and may be evaluated rapidly are devised and tested and we stress the advantages of using expressions which are linear in the variable parameters. For each system studied we have performed variational and diffusion quantum Monte Carlo calculations to test the accuracy of the optimized wave functions. The results of our study are very promising for future applications of quantum Monte Carlo methods to real materials.

I. INTRODUCTION

The optimization of approximate many-body wave functions is central to the development of quantum Monte Carlo (QMC) techniques. In the variational quantum Monte Carlo (VMC) method^{1,2} the energy is calculated as the expectation value of the Hamiltonian with an approximate wave function, the multidimensional integrals being evaluated via a Monte Carlo method. The variational principle guarantees that the energy obtained in the limit of an infinite VMC run (i.e., perfect integration) is an upper bound to the true ground state energy. In the diffusion quantum Monte Carlo (DMC) method,³⁻⁵ propagation of the imaginary-time Schrödinger equation is used to generate a probability distribution from which highly accurate energies can be calculated. Importance sampling is used, in which a guiding wave function is introduced to bias the sampling towards regions where the wave function is large, thereby reducing the variance of the estimate of the energy. In applications to large systems the fixed-node^{4,6} and short-time³ approximations are also used. The approximate wave function used in a VMC calculation (which also serves as the guiding wave function in a DMC calculation) normally contains variable parameters which may be optimized to improve its quality.

In both the VMC and DMC methods the role of the approximate/guiding wave function is crucial. The quality of the guiding wave function affects the accuracy of the DMC energy via the fixed-node approximation (i.e., how closely it approximates the nodal structure of the exact wave function).

In practice a more important issue for large systems is that the quality of the guiding wave function determines the variance of the estimate of the energy. The issue of statistical noise is particularly important in DMC calculations because the computational demands are roughly an order of magnitude more than for a VMC calculation using the same approximate/guiding wave function. Moreover, expectation values of operators which do not commute with the Hamiltonian are normally evaluated using the method of mixed estimation,³ which requires both the VMC and DMC expectation values and gives results which depend on the quality of the approximate/guiding wave function.

In this paper we report an investigation of wave function optimization using the method of minimization of the variance of the energy.⁷ We have studied three systems: the homogeneous electron gas (HEG), diamond-structure germanium with local pseudopotentials representing the ion cores, and the ground state of the corresponding germanium pseudoatom. Homogeneous systems are important in condensed matter theory because they are the simplest extended systems and they serve as a good testing ground for more realistic systems. We have developed a correlated wave function for the HEG which is accurate and allows for efficient optimization and rapid evaluation within a QMC calculation. We have used the electron correlation function developed for the HEG in calculations on diamond-structure germanium, adding a one-body term to introduce additional variational freedom in the charge density. Finally, we have studied the germanium pseudoatom. The issue of most interest here is

whether employing wave functions of similar flexibility in the pseudoatom and pseudosolid results in a large enough cancellation of errors to give a good value of the cohesive energy. For each of the three systems studied, the accuracy of the optimized wave functions is gauged by comparing with DMC results obtained using the optimized wave functions as guiding functions.

II. OPTIMIZATION BY MINIMIZATION OF THE VARIANCE OF THE ENERGY

One method of wave function optimization is to vary a set of parameters within the wave function so as to minimize the energy. Proper application of this method for a parametrized wave function gives the best (lowest) value for the energy of the system, but it may give poor values for other expectation values. If the energy is minimized then the local energy may be too high in some regions of configuration space and too low in others, so that the overall quality of the wave function is poor. This type of behavior contributes to the variance of the energy and therefore, minimization of the variance tends to give a better fit for the wave function as a whole, so that satisfactory results are obtained for a range of quantities, including the energy.⁷ Furthermore, the variance of the energy is zero for an eigenfunction and positive for an approximate wave function, and therefore the quantity to be minimized has a well-defined minimum value. The method of variance minimization was first applied to quantum mechanical problems in the 1930s. It was first used in QMC calculations by Coldwell,⁸ and some of the most impressive QMC applications have been by Umrigar *et al.*^{7,9}

We write the variance of the energy as

$$\sigma^2 = \sum_{\alpha} [\Psi_T^{-1}(R_{\alpha}) \hat{H} \Psi_T(R_{\alpha}) - \langle E_L \rangle]^2 \left(\frac{w(\alpha)}{\sum_{\beta} w(\beta)} \right), \quad (1)$$

where \hat{H} is the Hamiltonian and Ψ_T is the approximate/guiding wave function, which is to be varied. The sum is over the $3N$ -dimensional electron configurations, R_{α} , and $\langle E_L \rangle$ is the average energy,

$$\langle E_L \rangle = \sum_{\alpha} \Psi_T^{-1}(R_{\alpha}) \hat{H} \Psi_T(R_{\alpha}) \left(\frac{w(\alpha)}{\sum_{\beta} w(\beta)} \right), \quad (2)$$

where the reweighting factors, $w(\alpha)$, are given by

$$w(\alpha) = \left| \frac{\Psi_T(R_{\alpha})}{\Psi_T^0(R_{\alpha})} \right|^2. \quad (3)$$

The electron configurations are sampled from the starting distribution $|\Psi_T^0|^2$. This “correlated sampling” approach gives a good estimate of the difference in variance between wave functions corresponding to different sets of parameters. The process can be used iteratively by using the optimized set of parameters to regenerate configurations which are then used to perform a new optimization. This is useful when the reweighting factors differ significantly from unity. The nonlinear optimizations over the multidimensional parameter spaces were performed using the NAG routine E04FCF. This

works by finding the unconstrained minimum of a sum of squares, as in Eq. (1), using a modified Newton algorithm that requires the function values only.

III. APPROXIMATE/GUIDING WAVE FUNCTIONS FOR VMC AND DMC CALCULATIONS

The approximate wave functions used in VMC calculations and the guiding wave functions used in DMC calculations are usually chosen to be of the Slater-Jastrow form:

$$\Psi_T = D^{\uparrow} D^{\downarrow} \exp \left[- \sum_{i>j} u(\mathbf{r}_i, \mathbf{r}_j) + \sum_i \chi(\mathbf{r}_i) \right], \quad (4)$$

where D^{\uparrow} and D^{\downarrow} are Slater determinants of spin-up and spin-down single-particle orbitals, respectively. This form is not antisymmetric on exchange of electrons of opposite spin, but is convenient to use. Explicit antisymmetrization does not alter the expectation values of symmetric spin-independent operators such as the Hamiltonian, but would result in the replacement of two small determinants by a single large determinant, which would be computationally inefficient. The function $u(\mathbf{r}_i, \mathbf{r}_j)$ in the Jastrow factor correlates the motion of pairs of electrons, and $\chi(\mathbf{r}_i)$ is a one-body function. For each of the systems considered we have used somewhat different choices for u , χ , and D , which will be described below.

In previous QMC calculations a variety of different functional forms for u and χ have been used. For QMC calculations on solids the form for u has usually been restricted to functions depending only on the distance between electrons, a common form being a Yukawa function,¹⁰

$$u_{\sigma_i, \sigma_j}(r_{ij}) = \frac{A}{r_{ij}} \left[1 - \exp \left(- \frac{r_{ij}}{F_{\sigma_i, \sigma_j}} \right) \right], \quad (5)$$

where r_{ij} is the distance between electrons i and j , σ_i denotes the spin of electron i , and A and F_{σ_i, σ_j} are constants. In calculations using periodic boundary conditions the sum over electron pairs includes the infinite set of images of each electron. Normally the cusp conditions¹¹ are imposed which constrain the form of the wave function as $r_{ij} \rightarrow 0$, and lead to $F_{\sigma_i, \sigma_j} = \sqrt{2}A$ for parallel spins and $F_{\sigma_i, \sigma_j} = \sqrt{A}$ for anti-parallel spins. The only free parameter in u is then A which can be used as a variational parameter or set equal to $\sqrt{\Omega/4\pi N}$, where N/Ω is the number of electrons per unit volume. This choice for A is motivated by work on the homogeneous electron gas, which showed that at large electron separations u should have a $1/r_{ij}$ dependence with a coefficient of $1/\omega_p$, where ω_p is the plasma frequency.^{12,13} An alternative form for u in the HEG has been obtained within the random phase approximation.¹⁴

QMC calculations for light atoms and molecules are simpler than for solids, and therefore more extensive investigations of different correlation factors have been undertaken. The Boys-Handy correlation factor,¹⁵ which correlates the motion of two electrons with the position of a nearby nucleus, and variants thereof, have been used by several authors, and a similar correlation function has been applied to a solid by Mitás and Martin.¹⁶

In many QMC calculations, local density approximation (LDA) or Hartree-Fock (HF) orbitals have been used to construct the determinants in Eq. (4). A determinant of LDA or HF orbitals generally gives quite a good description of the charge density, but the addition of the correlation function, u , in Eq. (4) strongly modifies the charge density. The effect of u is to keep the electrons further from one another, thereby smoothing the charge density. The single-body χ function allows the electronic charge density to be altered in the presence of the u function.^{17,18} Fahy *et al.*^{17,18} used a form for χ which returned the charge density towards the LDA charge density. Fahy has recently developed a method for obtaining χ from u based on the random phase approximation which does not require an optimization and gives good results.¹⁹

IV. WAVE FUNCTION OPTIMIZATION FOR THE HOMOGENEOUS ELECTRON GAS

Recently we introduced a method for evaluating the expectation value of the Coulomb interaction in homogeneous systems which relies on the short range of the exchange-correlation hole.²⁰ This method allows accurate results to be obtained without summing the interactions over periodic images of the simulation cell, thereby reducing the computational cost significantly and leaving the evaluation of the Jastrow factor as the most time consuming part of the calculation. Therefore, we have sought a form for u which is accurate yet convenient for optimization purposes and can be evaluated rapidly within a QMC calculation.

The u function that we have developed is similar to one used earlier for the HEG by Ortiz and Ballone.²¹ In common with Ortiz and Ballone we use a spherically symmetric u function which is short ranged so that it need not be summed over simulation cells. This u function folds in the long-range behavior of the Jastrow factor in an approximate manner, and therefore depends on the size of the simulation cell, as well as on the electron density of the system. For each electron pair the separation vector \mathbf{r}_{ij} is reduced to its minimum length (by subtraction of supercell lattice vectors) giving the vector between electron i and the nearest periodic image of

electron j . The precise form for u that we have used is different from that used by Ortiz and Ballone, and our form has certain advantages which will be described below. We demand that u obeys the following conditions: (i) $u(r)$ satisfies the cusp conditions as $r \rightarrow 0$; (ii) $u(r)$ is continuous and has a continuous first derivative for all $r > 0$; (iii) $u(r)$ is linear in the variable parameters.

The only condition that $u(r)$ must satisfy for our QMC procedures to work is condition (ii) given above. If this condition is not obeyed then the kinetic energy estimator, $T_i = \frac{1}{4} \nabla_i^2 \ln \Psi$, will have δ functions at the discontinuities, which will be missed by the sampling procedure. To ensure continuity of the first derivative of $u(r)$ for $r > 0$ we require that u' goes (almost exactly) to zero at the surface of the sphere of radius L_{WS} inscribed within the Wigner-Seitz cell of the simulation cell. For $r > L_{WS}$ we set $u(r)$ and $u'(r)$ equal to zero. We normally impose the cusp conditions on the first derivative of u at $r \rightarrow 0$ because this is a property of the exact wave function. In contrast to Ortiz and Ballone we do not impose continuity of the second derivative of u . We write $u(r)$ as

$$u(r) = u_0(r) + f(r), \quad (6)$$

where u_0 is a fixed function and f contains the variable parameters. We expand f as a linear sum of some basis functions, f_k :

$$f(r) = \sum_k b_k f_k(r). \quad (7)$$

For the fixed part of u we choose

$$u_0(r) = \frac{A}{r} \left[1 - \exp\left(-\frac{r}{F}\right) \right] \exp\left(-\frac{r^2}{L_0^2}\right), \quad (8)$$

where F is chosen so that the cusp condition is obeyed and L_0 is chosen so that $u_0(L_{WS})$ is effectively zero ($< 10^{-6}$). Typically we choose $L_0 = 0.25 L_{WS}$ and A is fixed by the plasma frequency. The function u_0 is chosen to give a good description of the correlation so that the variable part of u is small. For the variable part we choose

$$f(r) = \begin{cases} B \left(\frac{L_{WS}}{2} + r \right) (L_{WS} - r)^2 + r^2 (L_{WS} - r)^2 \sum_{l=0}^M \alpha_l T_l(\bar{r}), & 0 \leq r \leq L_{WS} \\ 0, & r > L_{WS}, \end{cases} \quad (9)$$

where B and the α_l are variational coefficients, T_l is the l th Chebyshev polynomial, and

$$\bar{r} = \frac{2r - L_{WS}}{L_{WS}}, \quad (10)$$

so that the range $(0, L_{WS})$ is mapped into the orthogonality interval of the Chebyshev polynomials, $(-1, 1)$. The use of Chebyshev polynomials rather than a simple polynomial expression improves the numerical stability of the fitting pro-

cedure. The function f is the most general polynomial expression containing powers up to r^{M+4} which satisfies the following conditions: (i) $f'(0) = 0$; (ii) $f(r \geq L_{WS}) = 0$; (iii) $f'(r \geq L_{WS}) = 0$.

Condition (i) ensures that $u(r)$ obeys the cusp conditions, which are incorporated in $u_0(r)$. Addition of a constant to $u(r)$ changes the normalization of the wave function but not its functional form, and condition (ii) eliminates this unimportant degree of freedom. Condition (iii) ensures continuity of the first derivative of u at $r = L_{WS}$.

To start the optimization process we perform a VMC run to produce the electron configuration data for the initial distribution $|\Psi_T^0|^2$. For each electron configuration, $u(r)$ is summed over all distinct pairs of electron coordinates i and j in the simulation cell (with the separation vector reduced into the Wigner-Seitz simulation cell). For each configuration we write

$$\begin{aligned} \sum_{i>j} u(r_{ij}) &= \sum_{i>j} u_0(r_{ij}) + \sum_k b_k \sum_{i>j} f_k(r_{ij}) \\ &= \text{const} + \sum_k b_k g_k. \end{aligned} \quad (11)$$

Instead of storing the individual electron coordinates in each configuration we store the g_k , which is sufficient because our functional form for u is linear in the variable parameters. This reduces the storage and CPU time needed for the minimization procedure, which requires no further summations over the electron coordinates when the values of the parameters, b_k , are altered. The first and second derivatives of u , which enter the expression for the energy, are dealt with in a similar manner. These savings are very significant when dealing with a large number of electrons in the simulation cell, and for the HEG we have performed full minimizations with up to 338 electrons.

We have optimized wave functions for the HEG at a range of densities from $r_s=0.1$ to $r_s=10$. Excellent results were obtained at all densities, but for brevity we present results only for $r_s=1$. We used a wave function of Slater-Jastrow type [cf. Eq. (4)] where the determinants, $D^{\uparrow(\downarrow)}$, were constructed from the lowest energy plane waves at zero wave vector within the simulation cell Brillouin zone. The one-body χ function was set to zero and the u function of Eqs. (6)–(10) was used. Separate u functions for parallel and antiparallel spins were used, and we considered fcc simulation cells containing $N=30, 54, 178$, and 338 electrons. In each case the numbers of up- and down-spin electrons were equal. Typically we used 10 000 electron configurations sampled from a VMC run of sufficient length to ensure that the chosen configurations are statistically independent.

The variance minimization procedure is stable for small N , but gradually becomes unstable as N increases. The problem arises because the reweighting factors, $w(\alpha)$, in Eq. (1) deviate greatly from unity. Note that for a small change, δu , in the correlation factor, the local energy $(\Psi_T^{-1} \hat{H} \Psi_T)$ changes by an amount *proportional* to $N^2 \delta u$, but the reweighting factor is multiplied by a factor which is *exponential* in $N^2 \delta u$. For large N the behavior of the reweighting factor dominates the optimization process. One way to counteract this is to increase the number of configurations used, but this rapidly becomes impractical and we have found that it is better to reduce the importance of the reweighting factor. For $N \geq 54$ we set the reweighting factors equal to unity and generate several sets of configurations with successively better u functions until the procedure converges. This works because as the procedure converges the reweighting factors approach unity. At each iteration we allow only limited variations of the parameters which promotes the smooth convergence of the process. We have found this optimization procedure to be perfectly stable even up to the largest number of

electrons studied ($N=338$). Typically we use nine Chebyshev polynomials to represent $f(r)$, which tests show to give essentially complete convergence for the systems studied. The minimization problem then has 20 parameters.

In Table I we show the energy and standard deviation, σ , of the energy as a function of system size, comparing our u function [Eqs. (6)–(10)] with that of Eq. (5), which includes a sum over simulation cells, and with DMC results. Note that these values cannot be compared directly with those of other authors who use a different definition of the Coulomb energy of the finite simulation cell. In the limit of infinite system size these methods converge to the same energy, but for any finite size they are different, and the finite-size effects are smaller for our approach.²⁰ For the DMC calculations we used a time step of 0.01 a.u. and an average population of 640 configurations. After equilibration the averages were collected over 5000 moves of all the electrons. The results obtained using our new u function are of similar quality to those obtained with the u function of Eq. (5), but the new u function is much faster to evaluate. In Fig. 1(a) we show the optimized spin-parallel u function for $N=338$, together with the u function of Eq. (5) which is plotted in the [100] and [110] directions (for all other directions the u function lies between the values in these directions). In Fig. 1(b) the derivatives of the functions are shown. These figures show that the two functions are similar, but the optimized u function exhibits a slightly smaller derivative at intermediate distances. The optimized spin-antiparallel u function shows similar behavior.

The reduction in computing cost from using our new u function is very significant. It is particularly effective when combined with our recently developed technique for evaluating the expectation value of Coulomb interactions in homogeneous systems.²⁰ This combination of techniques entirely eliminates the need for time-consuming sums over simulation cells, and the resulting algorithm is extremely fast, with the most costly remaining operation being the calculation of determinants which are evaluated for each electron move using the standard Sherman-Morrison formula to update the matrix of cofactors.

V. A CRYSTALLINE SOLID: GERMANIUM

Having studied the HEG we now apply our newly developed u function to a crystalline solid. We have chosen to study germanium in the diamond structure, with a local pseudopotential of Starkloff-Joannopoulos form²² representing the Ge^{4+} ions. In previous works^{23,24} we performed VMC and DMC calculations on diamond-structure germanium using similar techniques to here, but variance minimization was not used in those calculations. In the present calculations we have achieved a considerably greater degree of optimization of the approximate/guiding wave functions than in our previous work. We have studied a fcc simulation cell of diamond structure germanium containing 16 atoms. The single particle orbitals were obtained from a LDA calculation using a plane-wave basis set containing all waves up to a kinetic energy cutoff of 40 Ry. The calculations were performed at a non-zero wave vector of the simulation cell Brillouin zone using the techniques described in Refs. 23 and 24. These techniques incorporate the ideas of “special k points”

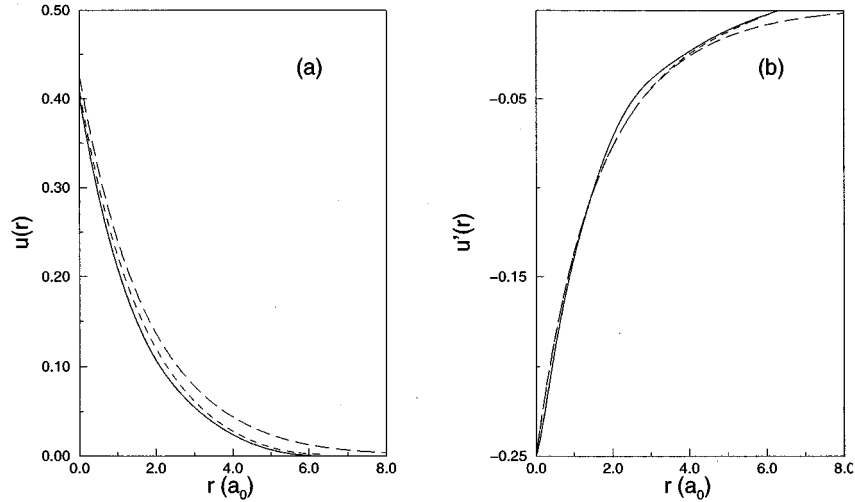


FIG. 1. Comparison of spin-parallel u functions for the HEG at $r_s = 1$. Our optimized function (solid line) is shown along with the Ewald summed Yukawa form along the (100) direction (long-dashed line) and the (110) direction (short-dashed line). (a) shows the u functions themselves while (b) shows the first derivatives.

which are used routinely in bandstructure calculations to improve the quality of the calculations.

The χ function was chosen to have the full symmetry of the diamond structure. We expand χ in a Fourier series:

$$\chi(\mathbf{r}) = \sum_{\mathbf{G}} c_{\mathbf{G}} e^{i\mathbf{G} \cdot \mathbf{r}}, \quad (12)$$

where $\{\mathbf{G}\}$ are the reciprocal lattice vectors of the primitive lattice. When performing the minimization we use the symmetry of the crystal and write the Fourier series as a sum over stars of reciprocal lattice vectors:

$$\chi(\mathbf{r}) = \sum_s c_s \left(\sum_{\mathbf{G} \in s} P_{\mathbf{G}} e^{i\mathbf{G} \cdot \mathbf{r}} \right), \quad (13)$$

where s labels a star of \mathbf{G} vectors and $P_{\mathbf{G}}$ is a phase factor associated with the non-symmorphic symmetry operations. In this expression each variable parameter, c_s , appears only once, so the minimization is unconstrained. We discard the $\mathbf{G} = 0$ term and also eliminate any terms whose coefficients are forced to be zero by the crystal symmetry. The Fourier series is linear in the variable parameters, so that we can perform the minimizations very efficiently as described in Sec. IV.

For the u function we used the functional form of Eqs. (6)–(10) which was developed for the HEG. The u and χ functions are optimized simultaneously because they are strongly coupled. Typically we use six nonzero coefficients in Eq. (13) for the χ function and eight parameters for both the parallel- and antiparallel-spin u functions in Eq. (9) giving a total of 22 parameters in the minimization problem. Variance minimizations were carried out using 10 000–100 000 independent N -electron configurations, which were regenerated several times. The final energy of -107.69 ± 0.01 eV per atom is 0.08 eV lower than the result obtained using the (Ewald summed) Yukawa potential of Eq. (5) and the variance minimization procedure for χ , and 0.20 eV lower than the result obtained in our previous work using the Yukawa potential and Fahy's original prescription for

χ .^{17,18} The energy of -107.69 eV per atom is only 0.34 eV per atom higher than the DMC result for this system of -108.03 ± 0.07 eV per atom quoted in Table I of Ref. 23. (As discussed in Refs. 23 and 24, we estimate that about 0.12 eV of this energy difference is due to the basis set incompleteness error in the single-particle orbitals, which affects the VMC much more than the DMC result, and which could be eliminated by the use of a larger basis set or a smoother pseudopotential.)

The optimized spin-parallel and spin-antiparallel u functions for germanium are similar to the Yukawa form in all directions. However, they have a smaller derivative at intermediate distances, exactly as observed in the HEG. The optimized χ function differs significantly from the original Fahy form, with some parameters changing by an order of magnitude. Altering the number of parameters in the optimization scheme revealed that six nonzero coefficients were sufficient to converge the χ function.

Recently we have performed similar optimization calculations using nonlocal pseudopotentials, which have yielded excellent results. We remark that the use of forms which are linear in the parameters is also advantageous in this case, as the extra integrations involved in using a nonlocal pseudopotential can be thought of as an increase in the number of electron configurations for this term in the energy.

In reality the electron correlation depends on the positions of all the electrons and all the atomic nuclei, rather than just on the distance between electrons, as we have so far assumed. As an attempt to include more complicated correlations we have investigated the effect of adding to the wave function a correlation function, u_s , which correlates the motion of two electrons with the position of a nearby atomic nucleus. The variables of our correlation function are the distances, r_1 and r_2 , of the electrons from the nucleus and the distance, r_{12} , between the electrons. This function will not help us describe the very short-ranged electron-electron correlations which lead to the cusp conditions and which depend only on the distance between the electrons, r_{12} , but instead we hope it will help us to describe the intermediate

range intra-atomic correlation effects. For the full correlation function we add to this the correlation function of Eqs. (6)–(10), which takes care of the cusp conditions and allows for long-range correlations. For the solid we sum u_s over the different atomic nuclei or ions, I , so that for each pair of electrons the correlation part of the wave function becomes

$$u(r_{12}) + \sum_I u_s(r_{1I}, r_{2I}, r_{12}). \quad (14)$$

The function $u_s(r_1, r_2, r_{12})$ should be symmetric under interchange of the electron coordinates r_1 and r_2 . Since u satisfies the cusp conditions, the function u_s should satisfy

$$\left. \frac{d}{dr_{12}} u_s(r_1, r_2, r_{12}) \right|_{r_{12}=0} = 0. \quad (15)$$

As u_s describes intermediate range correlations we cut it off smoothly at some distance L , which was varied between $3a_0$ and $5a_0$. Thus we impose

$$u_s(r_1, r_2, r_{12}) = 0, \quad r_1 \text{ or } r_2 \geq L \quad (16)$$

and

$$\left. \frac{d}{dr_1} u_s(r_1, r_2, r_{12}) \right|_{r_1=L} = 0, \quad r_1 \geq L, \quad (17)$$

and similarly for the r_2 derivative. We now write u_s as

$$u_s(r_1, r_2, r_{12}) = r_1^2 r_2^2 r_{12}^2 (L - r_1)^2 (L - r_2)^2 F(r_1, r_2, r_{12}), \quad (18)$$

where $F(r_1, r_2, r_{12})$ is symmetrical under interchange of r_1 and r_2 . F is expanded as a Chebyshev series:

$$F(r_1, r_2, r_{12}) = \sum_{i,j,k} \alpha_{ijk} T_i(\bar{r}_1) T_j(\bar{r}_2) T_k(\bar{r}_{12}), \quad (19)$$

where

$$\bar{r}_1 = \frac{2r_1 - L}{L}, \quad \bar{r}_2 = \frac{2r_2 - L}{L}, \quad \bar{r}_{12} = \frac{r_{12} - L}{L}, \quad (20)$$

and we enforce $\alpha_{ijk} = \alpha_{jik}$ to preserve the symmetry under interchange of electron coordinates. Note that an expansion of Eq. (18) about $r_1 = r_2 = r_{12} = 0$ gives the same first term as the Boys-Handy expression.

We have performed a variance minimization for diamond structure Ge using the u and u_s functions of Eqs. (6)–(10) and (18), and the χ function of Eq. (13). We used eight parameters for the spin-parallel and spin-antiparallel u functions, eight parameters for each of the u_s functions, and six parameters for the χ function, giving a total of 38 parameters in the wave function. The variance minimization was carried out over 100 000 independent configurations, which were regenerated several times. The final u_s function was very small and lowered the energy by only 0.01 eV per atom, and reduced the variance by only 1%. The type of correlations represented by u_s are therefore unimportant in this system.

VI. THE GERMANIUM PSEUDOATOM

The energy calculated within a VMC framework is an upper bound to the true energy of the system. When calculating energy differences it is important to use wave functions of similar quality for the different systems so that there is a strong cancellation of errors. This feature is common to many electronic structure methods. For instance, in band structure methods the errors due to incomplete basis sets are often large, but if sufficient care is taken it is often possible to calculate energy differences with high precision. A severe test of such cancellation is provided by the cohesive energy of a solid, which is the difference in energy between the solid and its constituent atoms. We have calculated the ground-state energy of the germanium pseudoatom using the same pseudopotential as for our solid calculations, and using a wave function of, as near as possible, the same quality as in the solid calculations.

The ground state of the germanium atom is a 3P configuration. The orbitals for the determinant were obtained from a LSDA calculation for the atomic ground state. We have used spherically symmetric χ functions for both up- and down-spin electrons of the form of the function f of Eq. (9) with $L_{\text{WS}} = 10a_0$. It is appropriate to choose this form for a pseudoatom because the electron-ion cusp condition implies that χ should be smooth at the origin. For the u functions we used the same form as for the HEG and germanium solid calculations [Eqs. (6)–(10)], with $L_{\text{WS}} = 10a_0$. We found that convergent results were obtained using eight parameters for both the χ and u functions. The energy obtained with the optimized wavefunction was -103.22 ± 0.01 eV and the standard deviation was 0.52 eV.

The χ functions for up- and down-spin electrons were found to be very similar. This result is perhaps surprising because there is a single down-spin electron in an s orbital, while there are three up-spin electrons, one in an s orbital and two in p orbitals. We investigated this point further by performing calculations using a single χ function, which gave an energy of -103.20 ± 0.01 eV, and a standard deviation of 0.54 eV, which are almost identical to the values obtained using separate functions for up- and down-spin electrons. The resulting χ and u functions were almost unchanged.

Figure 2 shows LDA and QMC charge densities for the pseudoatom. In Fig. 2(a) the VMC charge density from a wave function consisting of a determinant of LDA orbitals and a u function, but no χ function, is compared with the LDA density. This shows that the inclusion of the correlation factor, u , smears out the charge density considerably. In Fig. 2(b) we plot the charge density from an optimized wave function, containing both u and χ . This plot shows that the variance minimization procedure results in a wave function whose charge density is very close to the LDA form. In Fig. 2(c) we plot the DMC charge density calculated with the optimized wave function as guiding function. The DMC charge density is very close to both the VMC and LDA charge densities, which shows that the true charge density is close to the LDA form. It seems that the physical idea behind the original Fahy prescription for χ , i.e., returning the charge density to the LDA form, is extremely good. However, the number of configurations required to fit the variational pa-

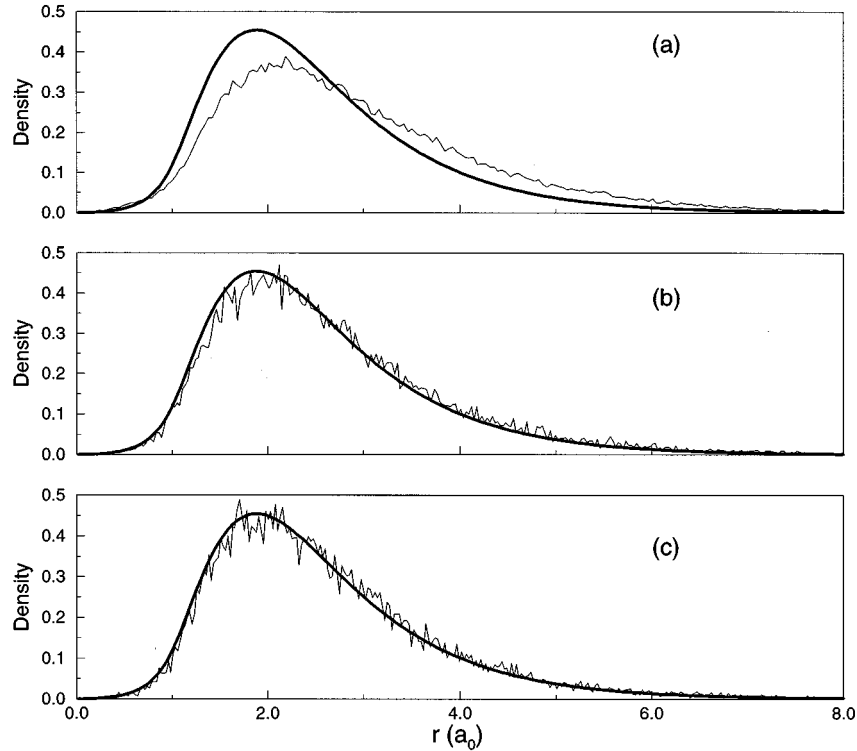


FIG. 2. Comparison of QMC densities (thin lines) and the LDA density (thick lines) of the germanium pseudoatom. (a) shows the VMC density calculated using a wavefunction consisting of a determinant of LDA orbitals and a u function, but with no χ function. (b) shows the VMC density calculated using a wave function containing optimized u and χ functions. (c) shows the DMC density calculated using the same optimized wave function as the guiding wave function.

rameters in a well-parametrized wave function via variance minimization is much less than the number required to obtain an accurate charge density. Therefore, it appears that for a given computational effort it is more efficient to generate χ via variance minimization than to construct the charge density and generate χ from it. Moreover, the method of variance minimization is more general in the sense that it does not rely on the separate determination of an accurate charge density from, for example, the LDA.

In order to compare the calculated cohesive energy with experiment, various correction terms must be added to the solid calculations to account for (i) Coulomb finite-size effects; (ii) single-particle finite-size effects; (iii) the use of a local pseudopotential; and (iv) zero-point motion. These corrections were discussed in detail in our previous work,²⁴ but we shall not be concerned with them here. They are essentially independent of the optimization of the wave function and it therefore suffices to compare directly the VMC and DMC results for the same system. A DMC calculation for the germanium pseudoatom gave an energy of -103.42 ± 0.03 eV, which is only 0.20 eV below our best VMC result, while for the solid the VMC calculation (without u_s) gave an energy 0.34 eV above the DMC result. Therefore, the VMC cohesive energy is only 0.14 eV less than the DMC result, which amounts to an error of only 4%. As mentioned earlier, we believe that 0.12 eV of the difference in energy between the VMC and DMC results for the solid is due to the incomplete basis set used for the single particle orbitals. If this is corrected for the VMC cohesive energy differs from the DMC value by only 1%.

VII. VARIANCE MINIMIZATION ON PARALLEL COMPUTERS

Most of the QMC calculations reported in this paper have been performed on parallel computers, and here we give a brief discussion of our implementation on such machines, concentrating on the variance minimization calculations. QMC calculations are “intrinsically parallel,” i.e., the QMC algorithm allows for a natural decomposition onto a set of parallel processors. For example, in VMC calculations we assign a fixed number of electron configurations to each processor, each of which undergoes an independent random walk, and interprocessor communication is needed only for the final collation of results. DMC calculations involve the creation and annihilation of electron configurations depending on global properties of the ensemble of configurations, so that more interprocessor communication is required. The number of electron configurations dealt with by each processor must be kept roughly equal during the simulation so as to maintain adequate load balancing. We will present more details of our implementation of VMC and DMC calculations on parallel computers in a future publication.²⁵

An important factor which determines the accuracy of the wave function obtained by variance minimization is the size of the ensemble of independent configurations over which one optimizes. We use large ensembles of up to 100 000 configurations containing several hundred electrons, and our wave functions contain roughly 20–40 variational parameters. The CPU and memory requirements are such that the advantages of using a parallel computer are considerable.

TABLE I. Energies, E , and standard deviations of the energies, σ , for the HEG at a density of $r_s = 1$ as a function of the number of electrons in the simulation cell, N . All entries are in Hartree atomic units per electron. The VMC energies are calculated with the Yukawa form of the Jastrow factor [Eq. (5)] and our optimized spherically symmetric form [Eqs. (6)–(10)]. The DMC energies do not depend on which of the two Jastrow factors is used.

N	E_{VMC} [Eq. (5)]	σ [Eq. (5)]	E_{VMC} [Eqs. (6)–(10)]	σ [Eqs. (6)–(10)]	E_{DMC}
30	0.4657	0.2	0.46979	0.22	
54	0.6085	0.24	0.6110	0.22	0.6069
178	0.6161	0.17	0.61679	0.17	0.6141
338	0.5772	0.14	0.57707	0.17	

With the current generation of parallel supercomputers, computational speeds at the multi-gigaflop level, together with gigabytes of distributed random access memory are standard. We have implemented the variance minimization algorithm described earlier on a 384-processor CRAY-T3D supercomputer at the Edinburgh Parallel Computing Centre. This machine is a multiple-instruction multiple-data (MIMD) computer. Each processor contains an identical copy of the program, but may execute different instructions on locally stored data. Communication between processors is carried out by message passing using the Parallel Virtual Machine (PVM) communication library.²⁶ Our implementation of the variance minimization procedure is based on the “master-slave” programming model where one processor, the master, delegates work to the other processors, the slaves. The master processor sends work to the slaves who complete the required work and return the results back to the master. The numerical optimization routine runs on the master processor and the ensemble of configurations is divided out among the slaves. The master processor broadcasts the values of the variational parameters to the slaves. Each of the slaves evaluates the reweighting factors, $w(\alpha)$, and the contributions to the variance [see Eq. (1)] for its subset of configurations. These are returned to the master which, via the NAG minimization routine, determines new values for the variational parameters. The procedure is repeated until the imposed limit on the change in the values of the parameters is reached, or until a minimum in the variance is found.

VIII. CONCLUSIONS

We have optimized wave functions for the homogeneous electron gas (HEG), a germanium pseudosolid, and the corresponding germanium pseudoatom, using the method of variance minimization. We have proposed forms for the Jastrow factor which are convenient to optimize and may be evaluated rapidly. Our strategy for the pair correlation term in the Jastrow factor is to use power expansions which are complete up to some given degree, and satisfy the various constraints imposed, while for the one-body function we use a Fourier expansion which satisfies the space group symmetry of the crystal. Both types of expression are amenable to systematic improvements by increasing the size of the basis set. We have stressed that using expressions which are linear in the variable parameters dramatically reduces the memory and CPU requirements of the optimization procedure. However, for the largest optimizations we have performed it is still advantageous to utilize the larger aggregate memory and

CPU power of a parallel computer. We have therefore implemented our optimization procedure on a distributed memory, MIMD, massively parallel supercomputer.

For the HEG we have performed optimizations with up to 338 electrons in the simulation cell. We find that for large numbers of electrons in the simulation cell the minimization procedure may become unstable if the reweighting factors deviate greatly from unity. A practical solution to this problem is to set the reweighting factors to unity and allow only small changes in the values of the parameters before regenerating another set of configurations and performing another minimization. For the HEG the energies obtained using our spherically symmetric u function are as good as those obtained with the common Yukawa form. However, our form does not require costly sums over simulation cells, and therefore it is computationally more efficient. This type of u function is particularly advantageous when used in conjunction with our formulation of the Coulomb interaction in homogeneous systems,²⁰ which obviates the need for sums over simulation cells in the interaction as well.

For the Ge solid, variance minimization applied to the one-body χ function is much more effective than the original Fahy prescription of choosing χ so as to return the charge density towards some desired form (normally the local-density approximation form). In this case we also found that our optimized polynomial expression for the u function gave superior results to the common Yukawa form. The best variational Monte Carlo (VMC) energy was only 0.34 eV per atom above the diffusion Monte Carlo (DMC) energy. We also tested a form of correlation factor similar to the Boys-Handy expression¹⁵ which correlates the motion of two electrons with the position of a nearby nucleus. Inclusion of this function resulted in a negligible reduction in the energy and its variance.

Wave function optimizations for the Ge pseudoatom in the 3P ground state configuration were performed using the same forms of correlation factor as in the solid and spherically symmetric χ functions for up- and down-spin electrons. Our best VMC energy was only 0.20 eV higher than the DMC result. The cohesive energy of Ge calculated using the VMC approach is only 0.14 eV smaller than the DMC result, which illustrates that it is possible to obtain comparable accuracy for the atom and solid using these optimization methods.

Overall, the results of our study are very promising for applications of the VMC technique to real materials, showing that for sp -bonded materials it should be possible to ob-

tain highly accurate results using fairly simple wave functions.

ACKNOWLEDGMENTS

We thank Mei-Yin Chou, Stephen Fahy and Cyrus Umrigar for helpful conversations. Financial support was provided by the Engineering and Physical Sciences Research Council (UK) under Grant Nos. GR/K21061 and GR/

H93064. Partial financial support was provided by the European Community under Contract No. CHRX CT94-0462 via the Human Capital and Mobility Research Network on Quantum Monte Carlo calculations for Solids. The computations were performed on the Cray T3D at the Edinburgh Parallel Computing Centre under Grant No. GR/K42318 and the Cray C-90 at the Pittsburgh Supercomputing Center under Grant No. DMR930039P.

*Author to whom correspondence should be addressed. Electronic address: ajw29@cam.ac.uk

¹W.L. McMillan, Phys. Rev. **138**, A442 (1965).

²D. Ceperley, G.V. Chester, and M.H. Kalos, Phys. Rev. B **16**, 3081 (1977).

³D.M. Ceperley and M.H. Kalos, in *Monte Carlo Methods in Statistical Physics*, edited by K. Binder (Springer-Verlag, Berlin, 1979); K.E. Schmidt and M.H. Kalos, in *Monte Carlo Methods in Statistical Physics II*, edited by K. Binder (Springer-Verlag, Berlin, 1984).

⁴D.M. Ceperley and B.J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

⁵P.J. Reynolds, D.M. Ceperley, B.J. Alder, and W.A. Lester, J. Chem. Phys. **77**, 5593 (1982).

⁶D.M. Ceperley, J. Stat. Phys. **63**, 1237 (1991).

⁷C.J. Umrigar, K.G. Wilson, and J.W. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988).

⁸R.L. Coldwell, Int. J. Quant. Chem. Symp. **11**, 215 (1977).

⁹C.J. Umrigar, M.P. Nightingale, and K.J. Runge, J. Chem. Phys. **99**, 2865 (1993).

¹⁰We use Hartree atomic units ($\hbar = e = m_e = 4\pi\epsilon_0 = 1$) for all equations.

¹¹T. Kato, Comm. Pure Appl. Math. **10**, 151 (1957).

¹²D.M. Ceperley, Phys. Rev. B **18**, 3126 (1978).

¹³D. Bohm and D. Pines, Phys. Rev. **92**, 609 (1953).

¹⁴T. Gaskell, Proc. Phys. Soc. **77**, 1182 (1961); **80**, 1091 (1962).

¹⁵S.F. Boys and N.C. Handy, Proc. R. Soc. London Ser. A **310**, 63 (1969).

¹⁶L. Mitas and R.M. Martin, Phys. Rev. Lett. **72**, 2438 (1994).

¹⁷S. Fahy, X.W. Wang, and S.G. Louie, Phys. Rev. Lett. **61**, 1631 (1988).

¹⁸S. Fahy, X.W. Wang, and S.G. Louie, Phys. Rev. B **42**, 3503 (1990).

¹⁹S. Fahy (unpublished).

²⁰L.M. Fraser, W.M.C. Foulkes, G. Rajagopal, R.J. Needs, S.D. Kenny, and A.J. Williamson, Phys. Rev. B **53**, 1814 (1996).

²¹G. Ortiz and P. Ballone, Phys. Rev. B **50**, 1391 (1994).

²²Th. Starkloff and J.D. Joannopoulos, Phys. Rev. B **16**, 5212 (1977).

²³G. Rajagopal, R.J. Needs, S. Kenny, W.M.C. Foulkes, and A. James, Phys. Rev. Lett. **73**, 1959 (1994).

²⁴G. Rajagopal, R.J. Needs, A. James, S. Kenny, and W.M.C. Foulkes, Phys. Rev. B **51**, 10 591 (1995).

²⁵A.J. Williamson, S.D. Kenny, G. Rajagopal, R.J. Needs, W.M.C. Foulkes, and P. Maccallum (unpublished).

²⁶A. Geist, A. Beguelin, J.J. Dongarra, W. Jiang, R. Manchek, and V.S. Sunderam, *PVM3 User's Guide and Reference Manual* (Oak Ridge National Laboratory, Oak Ridge, 1993).