PHYSICAL REVIEW B **66**, 155104 (2002)

Optimization of inhomogeneous electron correlation factors in periodic solids

David Prendergast and David Bevan*
Department of Physics, National University of Ireland, Cork, Ireland

Stephen Fahy
Department of Physics and National Microelectronics Research Centre, National University of Ireland, Cork, Ireland

(Received 8 February 2002; revised manuscript received 29 July 2002; published 9 October 2002)

A method is presented for the optimization of one-body and inhomogeneous two-body terms in correlated electronic wave functions of Jastrow-Slater type. The most general form of inhomogeneous correlation term which is compatible with crystal symmetry is used, and the energy is minimized with respect to all parameters using a rapidly convergent iterative approach, based on a Monte Carlo sampling of the energy and a fitting of energy fluctuations. The energy minimization is performed exactly within statistical sampling error for the energy derivatives, and the resulting one- and two-body terms of the wave function are found to be well determined. The largest calculations performed require the optimization of over 3000 parameters. The inhomogeneous two-electron correlation terms are calculated for diamond and rhombohedral graphite. The optimal terms in diamond are found to be approximately homogeneous and isotropic over all ranges of electron separation, but exhibit some inhomogeneity at short and intermediate ranges, whereas those in graphite are found to be homogeneous at short range, but inhomogeneous and anisotropic at intermediate- and long-range electron separations.

DOI: 10.1103/PhysRevB.66.155104 PACS number(s): 71.15.Nc, 02.70.Uu

I. INTRODUCTION

An accurate description of electron correlation is one of the central issues in modern electronic structure theory. In principle, this involves solving the many-electron Schrödinger equation, which for a system of N interacting electrons is an inseparable $3N$-dimensional problem. Hartree-Fock (HF) methods use a mean-field approximation of the electron interaction to reduce this to N independent problems, whose solution provides a reasonable approximation to many physical properties of the system. Density functional theory[1] (DFT) formally considers a noninteracting system with the same electronic density as the original, and is in principle an exact approach, although in practice requires approximations for its unknown exchange-correlation potential. While these approximations usually improve greatly on HF results, they may neglect some aspects of electron correlation which are chemically important or are associated with long-range electron correlation, such as van der Waals interactions.[2]

Quantum Monte Carlo (QMC) methods provide an important approach for solving the full $3N$-dimensional problem.[3] One such method, the variational Monte Carlo (VMC) method, allows us to estimate expectation values for a given trial wave function. Ideally, this wave function is an eigenstate of the many-body Hamiltonian. In practice, a parametrized form is used, which approximates the exact eigenstate. The accuracy of these calculations is entirely dependent on the trial wave function, however, and so the development of accurate wave functions is vital both in an accurate estimation of physical properties and in understanding how certain physical phenomena may be simply represented in the wave function.

A widely used trial correlated wave function is the Jastrow-Slater[4] or Feenberg[5] form

$$\Psi(\mathbf{r}_1,\ldots,\mathbf{r}_N)=\exp\left[-\sum_{i<j}u(\mathbf{r}_i,\mathbf{r}_j)+\sum_i\chi(\mathbf{r}_i)\right]D. \quad (1)$$

Here D is a Slater determinant of single-particle orbitals and interparticle correlation is introduced with the two-body term u in the Jastrow factor. The one-body term χ could in principle be absorbed into the single-particle orbitals of the determinant, but it may be convenient to retain it explicitly in the Jastrow factor. In practice, the orbitals in the Slater determinant are often determined from a HF or DFT calculation.[4]

In this paper, we will focus on the optimization of Jastrow-Slater wave functions in the context of the electronic structure of periodic solids. We will apply wave function optimization to examine some consequences which emerge from a complete treatment of inhomogeneity in the two-body correlation term for diamond and rhombohedral graphite.

The general form of wave function in Eq. (1) has been used as the starting point of several methods, including Fermi hypernetted chain[6] (FHNC) and VMC (Ref. 4) calculations. The traditional approach is to use a variational principle on the energy (or, in the ‘‘variance minimization’’ method,[7] on the fluctuations of the energy) to define a best approximation to the true eigenstate within the variational freedom allowed by the wave function ansatz. Expectation values of the energy and other quantities are calculated from the trial wave function, approximately in the FHNC approach, exactly in the VMC method (within statistical error of the sampling). While the VMC method has been used in calculations of periodic solids, in practice most calculations have included only homogeneous two-body terms in the Jastrow factor,[4] and the optimization of wave functions with very large numbers of parameters has remained problematical. To our knowledge, the FHNC approach has not been applied in fully three-dimensional electronic structure calculations.

Wave functions determined by the VMC approach are often used as guiding, trial functions for the diffusion Monte Carlo (DMC) method.[8] In the DMC context, the accuracy of the wave function affects the numerical efficiency of energy calculations and the accuracy of other physical quantities. When the DMC method is used in conjunction with nonlocal pseudopotentials,[9] as is commonly the case in chemical applications, an accurate trial wave function is essential for an accurate calculation of the ground state energy.

We present here a numerically robust, rapidly convergent iterative method which minimizes the variational energy with respect to a very general inhomogeneous form of the Jastrow factor, including all one- and two-body terms compatible with crystal symmetry. The method remains numerically well conditioned, even for large systems (of the order of several hundred electrons), where there are more than 3000 independent variational parameters in the Jastrow factor. Within acceptable computational demands on the Monte Carlo sampling, the optimal values of parameters in the wave function are found to be well determined, even when their contribution to the total energy is extremely small.

The method places no restrictions on the functional form of the antisymmetric (determinantal) part of the many-body wave function, and can be used in conjunction with related methods recently developed for energy minimization with respect to all orbitals in the determinant[10] or with respect to configuration weights in a multideterminant function.[11,12] Taken in conjunction with these methods, the approach presented here completes the solution to the problem of energy minimization with respect to the most general variational terms in wave functions of the Jastrow-Slater type, as well as in wave functions where the Jastrow factor multiplies a multideterminant function. Although many specific details of the work here refer to periodic systems, the basic method for Jastrow factor optimization could, in principle, be used in similar calculations of molecular, atomic, or nuclear structure.

Other general methods exist to achieve energy minimization with respect to parameters in many-body wave functions. When only one or two parameters are optimized, it is possible to perform a systematic search of parameter space, as McMillan did in his pioneering VMC study of the properties of liquid He.[13] The stochastic gradient approximation, used by Harju *et al.*,[14] applies control theory to determine iterative corrections to the wave function parameters. Lin *et al.*[15] explicitly computed analytical derivatives of the energy with respect to variational parameters, and used these to optimize their wave function with a Newton-style method. Snajdr *et al.*[16] used histogram filtering to optimize wave functions such that either the energy or its variance is minimized. However, none of these methods have as yet been applied in systems where a very large number of parameters are to be optimized.

In recent years, the variance minimization method of Umrigar *et al.*,[7] which optimizes the wave function by reducing the magnitude of the variance of the local energy, has been used much more widely than energy minimization in determining optimal parameters for variational wave functions. Although not strictly equivalent to energy minimization for a nonexact wave function, in practice this method has been very successful in improving total energies, providing extremely accurate wave functions in certain atomic systems.[7] However, recent studies for small systems (atoms and dimers) have shown that many non-energy-related properties of these systems are more accurately estimated by energy-optimized wave functions than by variance-optimized wave functions.[17]

Variance minimization may be thought of as fitting energy fluctuations to a constant, and attempting to reduce the cost function involved in this fit by a direct variation of the wave function parameters. The standard approach is to use a least-squares fit of these energy fluctuations.[44] In practice, the wave function parameters are often subject to confinement to local minima, and this approach requires much human interaction and experience for successful implementation when large numbers of parameters are to be optimized.

Our method of energy minimization involves the fitting of energy fluctuations to a given (nonconstant) functional form, which is a linear combination of operators associated with variations of the wave function parameters (see Sec. III). This fitting allows us to determine a ‘‘predictor,’’ which links changes in the wave function to changes in the fitted energy fluctuations. This predictor iteratively guides our method to a self-consistent solution, where the fitted energy fluctuations are zero. The derivatives of the true many-body energy with respect to all the parameters in the wave function are then also zero (within statistical sampling error) for this final solution.

Our predictor is closely related to the random phase approximation (RPA), introduced by Bohm and Pines[18] and recently discussed in the context of inhomogeneous systems.[19,20] However, as long as the predictor is sufficiently accurate to ensure a stable convergence of the iterations to the self-consistent solution, its exact form does not affect the final solution. Our method allows us to surpass the approximations of the RPA and to produce explicit trial wave functions of unprecedented accuracy for electrons in periodic solids.

We will use the optimized wave functions to study the effects of charge inhomogeneity on the correlation factors in diamond, as a prototype of strongly bonded insulating systems, and rhombohedral graphite, as a prototype of highly anisotropic, inhomogeneous solids. Earlier studies[4,21] of these and related systems indicated that homogeneous two-body correlation factors gain a large fraction of the correlation energy in solids. This seems to imply that there is little room for improvement by introducing inhomogeneous correlation factors. However, substantial differences in correlation factors often give rise to relatively little change in energy—a property that makes their optimization quite challenging. This is particularly true when one is interested in long-range correlation, which is energetically very delicate. Therefore, it is possible that substantial inhomogeneous structure may exist in these correlation factors, even though it contributes little to lowering the total energy of the system. We find this to be the case for graphite.

The rest of this paper is organized as follows: In Sec. II, we present the detailed numerical form of variational wave functions to be used in these calculations. The approach of fitting energy fluctuations and its use in guiding the iterative solution of the energy minimization problem is discussed in Sec. III. In Sec. IV, we present some results on the application of the method to the periodic solids, diamond and rhombohedral graphite, and examine the effects of charge density inhomogeneity on the correlation factors of these systems. In Sec. V we discuss the results and some computational details of the method, illustrating the justification for certain aspects of our approach with some tests. Finally, in Sec. VI, we present the overall conclusions of this study.

II. FORM OF THE WAVE FUNCTION

The Jastrow two-body correlation factor $u(\mathbf{r},\mathbf{r}')$ in Eq. (1) can in principle always be expressed as

$$u(\mathbf{r},\mathbf{r}')=\sum_{\alpha\beta} f_\alpha(\mathbf{r})^* u_{\alpha\beta} f_\beta(\mathbf{r}'), \tag{2}$$

where f_α form a complete set of functions and $u_{\alpha\beta}$ are expansion coefficients. (We use $*$ to indicate complex conjugation throughout this paper). Similarly, the one-body function $\chi(\mathbf{r})$ may also be expanded in a basis set of complete functions g_γ, as

$$\chi(\mathbf{r})=\sum_\gamma \chi_\gamma g_\gamma(\mathbf{r}). \tag{3}$$

Apart from an additional term to handle the electron-electron cusp,[22] we will express the Jastrow factor in the general form given by Eqs. (2) and (3). Values of the parameters $\{u_{\alpha\beta},\chi_\gamma\}$ will then be determined so that the energy of the system is stationary with respect to all variations $\{du_{\alpha\beta},d\chi_\gamma\}$.

Note that for the two-body function u, we shall retain its full inhomogeneity and anisotropy. By this, we mean that u is both a function of the position and relative orientation of two bodies positioned at $\mathbf{r}$ and $\mathbf{r}'$, i.e., $u=u(\mathbf{r},\mathbf{r}')$. This is in contrast to approximating u as: (i) a homogeneous function $u(\mathbf{r}'-\mathbf{r})$, dependent on relative position, or (ii) an isotropic function $u(|\mathbf{r}'-\mathbf{r}|)$, dependent only on the magnitude of separation, and so independent of orientation.

A. Electron-electron cusp

Due to the divergence of the Coulomb interaction between two electrons, the correct two-body correlation term $u(\mathbf{r},\mathbf{r}')$ has a cusp where $\mathbf{r}\to\mathbf{r}'$, leading to slow convergence of any expansion in smooth functions of the form in Eq. (2).[23] For this reason, it is numerically convenient to rewrite the two-body function u in the form

$$u(\mathbf{r},\mathbf{r}')+u_{\mathrm{sr}}(|\mathbf{r}-\mathbf{r}'|), \tag{4}$$

where $u_{\mathrm{sr}}(r)$ is a short-ranged, isotropic function which has the correct electron-electron cusp as $r\to 0$, and $u(\mathbf{r},\mathbf{r}')$ is now a smooth, cuspless function. We use a form of the short-ranged function u_{sr} which is generated from a numerical solution of the electron-electron scattering problem. A discussion of the generation of u_{sr} was provided in Appendixes A and C of Ref. 24. We define $u_{\mathrm{sr}}\equiv -\ln J_{\mathrm{sr}}$, where J_{sr} is defined within Appendix C of Ref. 24. The expansion of the remaining function $u(\mathbf{r},\mathbf{r}')$ in the form given in Eq. (2) then converges much more rapidly than that of the original function, for any set of smooth functions f_α.

In this paper, we generate the short-range function u_{sr} in a spin-dependent form, to maintain the cusp conditions.[22] The cuspless function $u(\mathbf{r},\mathbf{r}')$ used in our work is independent of the electron spin, although we expect that a spin-dependent form is as easily optimized.

B. Separation of one- and two-body terms

Summing $u(\mathbf{r},\mathbf{r}')$ over all electron pairs in the basis f_α leads to

$$\sum_{i<j} u(\mathbf{r}_i,\mathbf{r}_j)=\frac{1}{2}\sum_{\alpha\beta} u_{\alpha\beta}\sum_{i\neq j} f_\alpha(\mathbf{r}_i)^* f_\beta(\mathbf{r}_j). \tag{5}$$

This may be thought of as a two-body expansion of the correlation term. However, it is important to realize that any separation of "one-body" and "two-body" terms in the Jastrow factor is somewhat arbitrary. Removal of terms where $i=j$, as in Eq. (5), is not sufficient to decouple one- and two-body terms completely. To see this, consider the transformation of each of the basis functions obtained by subtracting a constant, $f'_\alpha(\mathbf{r})=f_\alpha(\mathbf{r})-c_\alpha$. The set remains complete and the function

$$u(\mathbf{r},\mathbf{r}')=\sum_{\alpha\beta}[f_\alpha(\mathbf{r})^*-c^*_\alpha]u_{\alpha\beta}[f_\beta(\mathbf{r}')-c_\beta] \tag{6}$$

may be interpreted as a "two-body" function in the basis set f'_α. However, expanding this correlation factor over all electron pairs, we find that, in terms of the original basis f_α, an additional one-body contribution appears:

$$\begin{aligned}\sum_{i<j} u(\mathbf{r}_i,\mathbf{r}_j)=&\frac{1}{2}\sum_{\alpha\beta} u_{\alpha\beta}\sum_{i\neq j} f_\alpha(\mathbf{r}_i)^* f_\beta(\mathbf{r}_j)\\ &-\frac{N-1}{2}\sum_{\alpha\beta}\sum_i [u_{\alpha\beta}f_\alpha(\mathbf{r}_i)^* c_\beta\\ &+u_{\alpha\beta}c^*_\alpha f_\beta(\mathbf{r}_i)]+\mathrm{const.}\end{aligned} \tag{7}$$

We may regard this as a transformation of the one-body function in Eq. (1), $\chi(\mathbf{r})\to\chi(\mathbf{r})+\chi^0(\mathbf{r})$, where the additional term comes from the second line of Eq. (7):

$$\begin{aligned}\chi^0(\mathbf{r})=&\frac{N-1}{2}\sum_\alpha\left\{\left[\sum_\beta u_{\alpha\beta}c_\beta\right]f_\alpha(\mathbf{r})^*\right.\\ &\left.+\left[\sum_\beta u_{\beta\alpha}c^*_\beta\right]f_\alpha(\mathbf{r})\right\}.\end{aligned} \tag{8}$$

To make a definite numerical separation of our one- and two-body expansions, we need to consider an appropriate choice of the arbitrary constants c_α, noting that $u(\mathbf{r},\mathbf{r}')$ is intended primarily to affect correlation properties (i.e., two-body properties) of the system, leaving single-particle properties unchanged. For example, the mean-field methods that produce D [HF, DFT under the local density approximation[25] (LDA), etc.] normally give very accurate single-particle densities, which can be altered substantially by inclusion of an arbitrary function $u(\mathbf{r},\mathbf{r}')$, in Ψ [Eq. (1)]. Any substantial change in the density from the HF solution is likely to be energetically very costly, so that ideally we would like to decouple changes in $u(\mathbf{r},\mathbf{r}')$ from changes in the density. Necessary changes in the density may be allowed for by optimization of the explicit one-body term χ in the Jastrow factor [Eq. (3)] or by the methods of Ref. 10.

Therefore, we would like to choose the constants c_α such that the average of any one-body operator (such as the density) for Ψ remains stationary with respect to variations in the coefficients $u_{\alpha\beta}$, at least in the absence of interparticle correlation. If we define the one-body operator

$$\rho_\gamma(\mathbf{R}) \equiv \sum_i f_\gamma(\mathbf{r}_i)^*$$

for the many-body configuration $\mathbf{R}=(\mathbf{r}_1 \ldots \mathbf{r}_N)$ then its expectation value is

$$\langle \rho_\gamma \rangle = \langle \Psi | \rho_\gamma | \Psi \rangle = \int f_\gamma(\mathbf{r})^* \rho(\mathbf{r}) d\mathbf{r},$$

where $\rho(\mathbf{r})$ is the single particle density. Since the f_γ are fixed functions, variations in the expectation values $\langle \rho_\gamma \rangle$ correspond to variations in $\rho(\mathbf{r})$. The derivative of $\langle \rho_\gamma \rangle$ with respect to variations of the $u_{\alpha\beta}$ in Eq. (6) is (see Appendix A)

$$\begin{aligned}\frac{1}{2}\frac{\partial\langle\rho_\gamma\rangle}{\partial u_{\alpha\beta}} &= \Big\langle [\rho_\gamma(\mathbf{R}) - \langle\rho_\gamma\rangle] \\ &\quad \times \sum_{i\neq j} [f_\alpha(\mathbf{r}_i)^* - c_\alpha^*][f_\beta(\mathbf{r}_j) - c_\beta] \Big\rangle \\ &= \sum_k \sum_{i\neq j} \langle [f_\gamma(\mathbf{r}_k)^* - \langle f_\gamma\rangle^*] \\ &\quad \times [f_\alpha(\mathbf{r}_i)^* - c_\alpha^*][f_\beta(\mathbf{r}_j) - c_\beta] \rangle.\end{aligned}$$

In the absence of interparticle correlation,

$$\langle f_\gamma(\mathbf{r}_k)^* f_\alpha(\mathbf{r}_i)^* f_\beta(\mathbf{r}_j) \rangle = \begin{cases} \langle f_\gamma^* f_\alpha^* \rangle \langle f_\beta \rangle & \text{if } k=i \\ \langle f_\gamma^* f_\beta \rangle \langle f_\alpha \rangle^* & \text{if } k=j \\ \langle f_\gamma \rangle^* \langle f_\alpha \rangle^* \langle f_\beta \rangle & \text{otherwise.} \end{cases}$$

The second summation excludes $i=j$, so averages of triple products never arise. Thus, in the absence of correlation, the derivative becomes

$$\begin{aligned}\frac{1}{2}\frac{\partial\langle\rho_\gamma\rangle}{\partial u_{\alpha\beta}} &\approx \sum_{i\neq j} \langle [f_\gamma(\mathbf{r}_i)^* - \langle f_\gamma\rangle^*][f_\alpha(\mathbf{r}_i)^* - c_\alpha^*] \rangle \\ &\quad \times \langle f_\beta(\mathbf{r}_j) - c_\beta \rangle + \langle [f_\gamma(\mathbf{r}_j)^* - \langle f_\gamma\rangle^*] \\ &\quad \times [f_\beta(\mathbf{r}_j) - c_\beta] \rangle \langle f_\alpha(\mathbf{r}_i)^* - c_\alpha^* \rangle. \qquad (9)\end{aligned}$$

We can guarantee that the right hand side of Eq. (9) is zero if $c_\alpha = \langle f_\alpha \rangle$, for all α. In other words, one-body expectation values remain approximately unaffected by the presence of the correlation factor $u(\mathbf{r},\mathbf{r}')$, provided we expand $u(\mathbf{r},\mathbf{r}')$ in a basis of "fluctuation functions," $f'_\alpha(\mathbf{r}) = f_\alpha(\mathbf{r}) - \langle f_\alpha \rangle$. Equivalently, we may retain the original basis f_α and form a one-body term χ^0 from Eq. (8) with $c_\alpha = \langle f_\alpha \rangle$. This may then be inserted into the wave function of Eq. (1) and varied as the parameters $u_{\alpha\beta}$ are varied.

Correlation effects are of course present in the actual wave function used. However, we find that Eq. (9) remains approximately true, as previously observed.[21,26] For the energy minimization problem, we find that mixed derivatives of the energy $\partial^2 E/\partial u_{\alpha\beta}\partial\chi_\gamma$ are approximately zero when $c_\alpha = \langle f_\alpha \rangle$. This gives the numerical advantage that minimization of the energy with respect to $u_{\alpha\beta}$ approximately decouples from minimization with respect to χ_γ. We note that this approximate decoupling holds, even if the relation $c_\alpha = \langle f_\alpha \rangle$ is not exactly true. Thus we may use $c_\alpha = \langle D | f_\alpha | D \rangle$, (i.e., LDA or HF averages of f_α), in place of $\langle \Psi | f_\alpha | \Psi \rangle$, while still maintaining the numerical advantages of approximately satisfying $\partial^2 E/\partial u_{\alpha\beta}\partial\chi_\gamma = 0$.

C. Fourier expansion

In the context of periodic systems, it is natural to expand the correlation function $u(\mathbf{r},\mathbf{r}')$ as a Fourier series, where the basis functions are $f_\mathbf{q} = \exp[i\mathbf{q}\cdot\mathbf{r}]$, for each wave vector $\mathbf{q}$. We note that

$$\rho_\mathbf{q}(\mathbf{R}) = \sum_i \exp[-i\mathbf{q}\cdot\mathbf{r}_i] = \sum_i f_\mathbf{q}(\mathbf{r}_i)^* \qquad (10)$$

is the Fourier coefficient of the instantaneous charge density, given the electron configuration $\mathbf{R}$. (In atomic units the electronic charge $e=1$). Also, summing over pairs leads to a quadratic product of Fourier coefficients, with some modification to remove terms with $i=j$:

$$\sum_{i<j} f_\mathbf{q}(\mathbf{r}_i)^* f_{\mathbf{q}'}(\mathbf{r}_j) \equiv \frac{1}{2} (\rho_\mathbf{q} \rho_{\mathbf{q}'}^*)_{[i\neq j]}. \qquad (11)$$

In order to approximately remove the effect of the two-body terms on the single-particle density, following Sec. II B, we subtract appropriate constants from each basis function to produce a basis of collective "charge fluctuation" coordinates

$$\sum_i f'_\mathbf{q}(\mathbf{r}_i)^* \equiv \Delta\rho_\mathbf{q} = \rho_\mathbf{q} - \langle \rho_\mathbf{q} \rangle = \sum_i f_\mathbf{q}^*(\mathbf{r}_i) - \langle f_\mathbf{q} \rangle^*. \qquad (12)$$

These provide a suitable expansion of the two-body correlation factor that approximately preserves single-particle densities.

A correlation factor using such coordinates was first suggested by Bohm and Pines[18] for the homogeneous electron gas, and has recently been discussed by Malatesta *et al.*[21] and Gaudoin *et al.*[20] in the context of inhomogeneous systems. In homogeneous systems the expectation value $\langle \rho_{\mathbf{q}} \rangle$ of each charge density Fourier coefficient is zero, and so charge density fluctuations are simply $\rho_{\mathbf{q}}$. For inhomogeneous systems, in general $\langle \rho_{\mathbf{q}} \rangle \neq 0$ when $\mathbf{q}=\mathbf{G}$, a reciprocal lattice vector.

The alternative to using fluctuation coordinates $\Delta \rho_{\mathbf{q}}$ is to incorporate the equivalent one-body term in the Jastrow factor, which in periodic systems is of the form

$$\sum_i \chi^0(\mathbf{r}_i) = \sum_{\mathbf{G}} \chi^0_{\mathbf{G}} \rho^*_{\mathbf{G}},$$

with the coefficients coming from the two-body term

$$\chi^0_{\mathbf{G}} = \frac{N-1}{N} \sum_{\mathbf{G}'} u_{\mathbf{G}\mathbf{G}'} \langle \rho_{\mathbf{G}'} \rangle, \tag{13}$$

as discussed by Malatesta *et al.*[21] and Gaudoin *et al.*[20]

The properties of the Fourier basis $f_{\mathbf{q}}$ and the correlation factor u lead to some convenient symmetry properties for the coefficients $u_{\mathbf{q}\mathbf{q}'}$. The complex conjugate $u^*_{\mathbf{q}\mathbf{q}'} = u_{-\mathbf{q}-\mathbf{q}'}$, just as $f_{\mathbf{q}} = f^*_{-\mathbf{q}}$. The exchange symmetry of u, i.e., $u(\mathbf{r},\mathbf{r}') = u(\mathbf{r}',\mathbf{r})$ implies that $u_{\mathbf{q}\mathbf{q}'} = u_{-\mathbf{q}'-\mathbf{q}}$. If u possesses inversion symmetry, i.e., $u(\mathbf{r},\mathbf{r}') = u(-\mathbf{r},-\mathbf{r}')$, then each $u_{\mathbf{q}\mathbf{q}'}$ is a real number. In periodic systems,

$$u(\mathbf{r}+\mathbf{L},\mathbf{r}'+\mathbf{L}) = u(\mathbf{r},\mathbf{r}')$$

for any Bravais lattice vector $\mathbf{L}$. This implies that all Fourier coefficients $u_{\mathbf{q}\mathbf{q}'}$ are zero unless $\mathbf{q}-\mathbf{q}'=\mathbf{G}$, a reciprocal lattice vector. Thus translation symmetry greatly reduces the number of variational parameters in the two-body terms.

We arrange the wave function in the form

$$\Psi = J_{\mathrm{sr}} J_{\mathrm{ih}} J_{1\mathrm{b}} D,$$

isolating the short-range component of the Jastrow factor as

$$J_{\mathrm{sr}} = \exp\left[\sum_{|\mathbf{G}|<G_c} \chi^0_{\mathrm{sr}}(\mathbf{G}) \rho^*_{\mathbf{G}} - \sum_{i<j} u_{\mathrm{sr}}(r_{ij}) \right].$$

The one-body term here is derived from the short-range correlation factor u_{sr} of Eq. (4), as $\chi^0_{\mathrm{sr}}(\mathbf{G}) = \tilde{u}_{\mathrm{sr}}(\mathbf{G}) \langle \rho_{\mathbf{G}} \rangle$, where $\tilde{u}_{\mathrm{sr}}(\mathbf{G})$ is the Fourier transform of u_{sr} for the reciprocal lattice vector $\mathbf{G}$. The prefactor of $(N-1)/N$, which should be present from Eq. (13), approaches unity for large systems, and so is neglected. For computational efficiency we leave this one-body term in its Fourier space representation, and G_c is a cutoff chosen for the Fourier sum such that it is converged within a required accuracy.

The remaining inhomogeneous part of the two-body Jastrow factor is expanded using charge density fluctuation coordinates

$$J_{\mathrm{ih}} = \exp\left[-\sum_{\mathbf{q}} \sum_{\mathbf{G}\mathbf{G}'} u_{\mathbf{q}+\mathbf{G}\mathbf{q}+\mathbf{G}'} P_{\mathbf{q}+\mathbf{G}\mathbf{q}+\mathbf{G}'} \right], \tag{14}$$

where $P_{\mathbf{q}+\mathbf{G}\mathbf{q}+\mathbf{G}'} \equiv (\Delta \rho_{\mathbf{q}+\mathbf{G}} \Delta \rho_{\mathbf{q}+\mathbf{G}'})_{[i\neq j]}$, using the notation $(\cdot)_{[i\neq j]}$, as defined in Eq. (11) and the definition of $\Delta \rho_{\mathbf{q}+\mathbf{G}}$ in Eq. (12). In practice, this double sum is truncated by using vectors $\mathbf{q}+\mathbf{G}$ of magnitude less than a suitably chosen cutoff k_c.

We also allow for one-body optimization through the use of the explicit one-body Jastrow factor. This one-body Jastrow factor is also expanded in fluctuation coordinates,

$$J_{1\mathrm{b}} = \exp\left[\sum_{|\mathbf{G}|<G_c} \chi_{\mathbf{G}} \Delta \rho^*_{\mathbf{G}} \right], \tag{15}$$

since including the constant average values $\langle \rho_{\mathbf{G}} \rangle$ merely adjusts the normalization of the wave function.

The coefficients $u_{\mathbf{q}+\mathbf{G}\mathbf{q}+\mathbf{G}'}$ and $\chi_{\mathbf{G}}$, defined in Eqs. (14) and (15), are the final variational parameters of our wave function. The remaining sections of this paper describe the method we use to optimize these parameters such that the total energy of a particular electronic system is stationary. A typical calculation presented below involves the simultaneous optimization of over 3000 parameters.

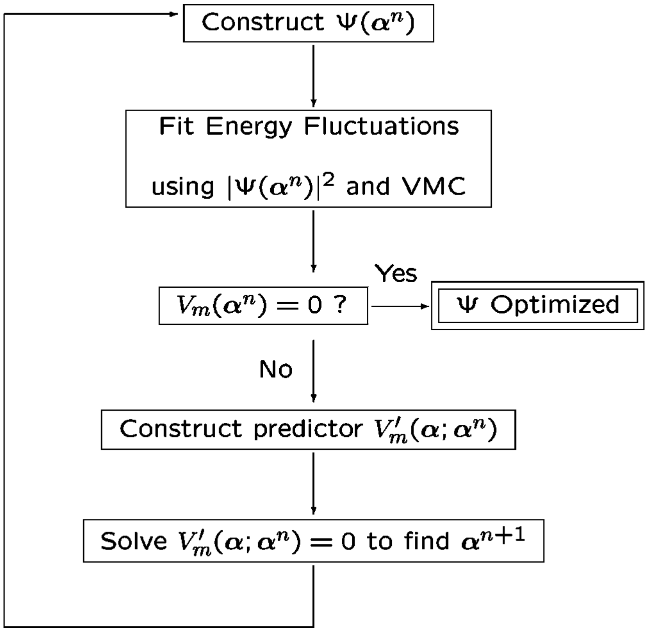

FIG. 1. A flow chart outlining the iterative method of wave function optimization by a minimization of the total energy of the system, as described in Sec. III.

III. ENERGY MINIMIZATION

This section describes in detail an iterative approach to solving the energy minimization problem for a given parametrized wave function. The various steps involved in the method are presented in the form of a flow chart in Fig. 1, and are described now in brief (precise definitions will be provided in subsequent sections for the fitting coefficients and predictor function mentioned here).

(1) For a given set of variational parameters $\boldsymbol{\alpha}^n$ we may construct the numerical trial wave function $\Psi(\boldsymbol{\alpha}^n)$. (2) Using this wave function to guide a VMC calculation, we perform a least-squares fit of the fluctuations in the energy of the system to a prescribed functional form, producing a set of fitting coefficients $V_m(\boldsymbol{\alpha}^n)$, which are themselves functions of the wave function parameters $\boldsymbol{\alpha}^n$. (3) If these V_m are zero, then the wave function is optimized with respect to its parameters. (4) However, if they are not zero, then we must attempt to set them to zero. This is not straightforward, since we do not know the exact functional dependence of V_m on $\boldsymbol{\alpha}$, and so, we construct our best estimate of this dependence—the predictor $V'_m(\boldsymbol{\alpha},\boldsymbol{\alpha}^n)$—defined for the given set of parameters $\boldsymbol{\alpha}^n$. (5) We find the set of parameters $\boldsymbol{\alpha}^{n+1}$ that sets this predictor to zero and use this as the next input for construction of the wave function in step (1). This process continues until the V_m are zero and the wave function is optimized.

A. Euler-Lagrange equations

We wish to optimize the wave function

$$\Psi=\Psi(\boldsymbol{\alpha}),$$

where $\boldsymbol{\alpha}=\{\alpha_m\}$ a vector of parameters, by solving the Euler-Lagrange equations

$$\frac{\partial\langle\mathcal{H}\rangle}{\partial\alpha_m}=0, \quad \text{for all m.} \tag{16}$$

We note that the Hamiltonian for the system is

$$\mathcal{H}=-\frac{1}{2}\sum_i \nabla_i^2+\sum_i V_{\mathrm{ext}}(\mathbf{r}_i)+\sum_{i<j} V(r_{ij}), \tag{17}$$

where each sum is over the electrons in the system.

Solving Eq. (16) is equivalent (see Appendix A) to solving the system of equations

$$\langle\Delta\mathcal{H}\Delta\mathcal{O}_m\rangle=0 \quad \text{for all } m, \tag{18}$$

where, given a many-body configuration $\mathbf{R}=\{\mathbf{r}_i\}$, we define $\Delta A(\mathbf{R})\equiv A(\mathbf{R})-\langle A\rangle$ for any operator A; and the local values of the operators $\mathcal{O}_m$ as,

$$\mathcal{O}_m(\mathbf{R})\equiv\frac{\partial}{\partial\alpha_m}\ln\Psi(\mathbf{R})=\begin{cases}-P_{\mathbf{qq}'}(\mathbf{R}) & \text{for} \quad \alpha_m=u_{\mathbf{qq}'}\\ \Delta\rho_{\mathbf{G}}(\mathbf{R})^* & \text{for} \quad \alpha_m=\chi_{\mathbf{G}}.\end{cases} \tag{19}$$

We shall refer to the local value of the Hamiltonian operator as the "local energy":

$$E(\mathbf{R})\equiv\frac{\mathcal{H}\Psi(\mathbf{R})}{\Psi(\mathbf{R})}. \tag{20}$$

We approach the problem of solving the Euler-Lagrange equations [Eq. (16)] indirectly, by considering systematic fluctuations of the energy for a given trial wave function $\Psi(\boldsymbol{\alpha})$.

B. Systematic energy fluctuations

Consider fitting the local energy $E(\mathbf{R})$ to the functional form

$$E_0+\sum_m V_m\mathcal{O}_m(\mathbf{R}), \tag{21}$$

in the least-squares sense, where $\{\mathcal{O}_m\}$ is the set of functions with which we fit the energy, and $\{V_m\}$ is the vector of fitting coefficients. The least-squares problem reduces to minimizing the integral

$$\langle\Psi|\left\{\mathcal{H}-E_0-\sum_m V_m\mathcal{O}_m\right\}^2|\Psi\rangle,$$

which is equivalent (see Appendix B) to solving the linear system

$$\sum_m V_m\langle\Delta\mathcal{O}_m\Delta\mathcal{O}_n\rangle=\langle\Delta E\Delta\mathcal{O}_n\rangle \quad \text{for all } n. \tag{22}$$

We recognize immediately that if the functions $\mathcal{O}_m$ are those functions associated with variations of the wave function parameters α_m [Eq. (19)], then the right-hand side of Eq. (22) *is* the vector of Euler-Lagrange derivatives in Eq. (18). Therefore, the Euler-Lagrange equations [Eq. (16)] are solved if all the fitted coefficients V_m are zero.[45]

As an illustrative example, consider Ψ_0, an eigenstate of $\mathcal{H}$. Indeed, the local energy is a constant, independent of $\mathbf{R}$, and so we would find that each of the fitted coefficients V_m is zero. Therefore, the Euler-Lagrange derivatives $\langle\Delta\mathcal{H}\Delta\mathcal{O}_m\rangle$ are all zero, and so the energy must be stationary with respect to variations in Ψ_0, as we would expect for an eigenstate.

For the trial wave function $\Psi(\boldsymbol{\alpha})$, no choice of $\boldsymbol{\alpha}$ gives an *exact* eigenstate of the Hamiltonian. However, for a particular choice of parameters, the absence of *systematic* variations of the energy (i.e., variations correlated with the variations of the functions $\mathcal{O}_m$) ensures that the fitting coefficients V_m are all zero and that the average energy is stationary with respect to all variations of the parameters $\boldsymbol{\alpha}$. Within the parametric freedom of the trial wave function Ψ, this is our best approximation to an eigenstate.

C. Iterative procedure

We now describe a procedure which aims, by appropriate choice of the parameters $\boldsymbol{\alpha}$, to set the fitted coefficients V_m of the total energy E to zero. As defined in Eq. (22), these V_m depend on the wave function $\Psi(\boldsymbol{\alpha})$ and so are functions of the parameters $\boldsymbol{\alpha}$. However, the functional dependence of V_m on $\boldsymbol{\alpha}$ is not available in its exact analytic form, and we are unable to solve directly the system $V_m(\boldsymbol{\alpha})=0$, i.e., to find the root $\boldsymbol{\alpha}$ which will guarantee the solution to the corresponding Euler-Lagrange equations.

Instead, using the wave function $\Psi(\boldsymbol{\alpha}^0)$, for a particular choice of the parameters $\boldsymbol{\alpha}^0$, we construct a predictor function $V'_m(\boldsymbol{\alpha};\boldsymbol{\alpha}^0)$, which approximates the unknown function $V_m(\boldsymbol{\alpha})$ for general values of $\boldsymbol{\alpha}$. More precisely, we construct $V'_m(\boldsymbol{\alpha};\boldsymbol{\alpha}^0)$ so that

$$V'_m(\boldsymbol{\alpha}^0;\boldsymbol{\alpha}^0)\equiv V_m(\boldsymbol{\alpha}^0) \tag{23}$$

(i.e., V'_m is exact when $\boldsymbol{\alpha}=\boldsymbol{\alpha}^0$) and

$$V'_m(\boldsymbol{\alpha};\boldsymbol{\alpha}^0)\approx V_m(\boldsymbol{\alpha}) \tag{24}$$

for all relevant values of $\boldsymbol{\alpha}$.

To determine this predictor, we use the specific form of the Hamiltonian [Eq. (17)] and trial wave function Ψ (Sec. II), and we partition the local energy $E(\mathbf{R})$ into a sum of contributions,

$$E(\mathbf{R})=\sum_i \epsilon^{(i)}(\mathbf{R}), \tag{25}$$

where the $\epsilon^{(i)}$ come from various terms in the kinetic and potential energy (see below). Each contribution $\epsilon^{(i)}(\mathbf{R})$ is approximated with the functional form

$$\epsilon_0^{(i)}+\sum_m v_m^{(i)}\mathcal{O}_m(\mathbf{R}). \tag{26}$$

For some terms, using the specific form of the local energy for $\Psi(\boldsymbol{\alpha})$, we can expand $\epsilon^{(i)}$ analytically as in Eq. (26), enabling us to determine, exactly or approximately, the function $v_m^{(i)}(\boldsymbol{\alpha})$. In this analytic form, $v_m^{(i)}(\boldsymbol{\alpha})$ is independent of the choice of $\boldsymbol{\alpha}^0$ and so remains equally valid for all $\boldsymbol{\alpha}$.

Where analytic expressions are too complex to derive, we may approximate $v_m^{(i)}(\boldsymbol{\alpha})$ by fitting $\epsilon^{(i)}$ to Eq. (26). The fitting coefficients for $\epsilon^{(i)}$ are found by solving the analog of Eq. (22),

$$\sum_m v_m^{(i)}\langle\Delta\mathcal{O}_m\Delta\mathcal{O}_n\rangle=\langle\Delta\epsilon^{(i)}\Delta\mathcal{O}_n\rangle \quad \text{for all } n, \tag{27}$$

where $v_m^{(i)}=v_m^{(i)}(\boldsymbol{\alpha}^0)$ is determined by using $\Psi(\boldsymbol{\alpha}^0)$ to evaluate the required expectation values. This produces the value of each coefficient at $\boldsymbol{\alpha}^0$. We may also fit the derivatives of $\epsilon^{(i)}$ with respect to $\boldsymbol{\alpha}$, to determine the linear dependence of each $v_m^{(i)}(\boldsymbol{\alpha})$. We may then approximate the function

$$v_m^{(i)}(\boldsymbol{\alpha})\approx v_m^{(i)}(\boldsymbol{\alpha};\boldsymbol{\alpha}^0)=v_m^{(i)}(\boldsymbol{\alpha}^0)+\sum_l \left.\frac{\partial v_m^{(i)}}{\partial\alpha_l}\right|_{\boldsymbol{\alpha}^0}(\alpha_l-\alpha_l^0), \tag{28}$$

where the values of the derivatives are found by fitting $\partial\epsilon^{(i)}/\partial\alpha_l$ to Eq. (26) using $\Psi(\boldsymbol{\alpha}^0)$. In evaluating the term $\partial v_m^{(i)}/\partial\alpha_l$, we consider only the explicit variation of the term $v_m^{(i)}$ with α_l. We do not include the implicit variation due to the dependence of the probability distribution $|\Psi(\boldsymbol{\alpha})|^2$ on $\boldsymbol{\alpha}$. In practice, the only term for which we need to fit $\partial\epsilon^{(i)}/\partial\alpha_l$ is explicitly linear in the parameters $\boldsymbol{\alpha}$ and so the linear expansion in Eq. (28) is valid over a wide range of values of $\boldsymbol{\alpha}$.

Just as the energy contributions $\epsilon^{(i)}(\mathbf{R})$ partition the local energy, we may regard the analytic and fitted coefficients $v_m^{(i)}$ as an approximate partition of the local energy coefficients V_m. We define this partition as

$$W_m(\boldsymbol{\alpha};\boldsymbol{\alpha}^0)\equiv S_m(\boldsymbol{\alpha})+T_m(\boldsymbol{\alpha};\boldsymbol{\alpha}^0),$$

where the sum of analytically derived coefficients is

$$S_m(\boldsymbol{\alpha})\equiv\sum_{\text{analytic}} v_m^{(i)}(\boldsymbol{\alpha}),$$

and the sum of numerically determined coefficients, evaluated using $\Psi(\boldsymbol{\alpha}^0)$ and Eq. (27), is

$$T_m(\boldsymbol{\alpha};\boldsymbol{\alpha}^0)\equiv\sum_{\text{fitted}} v_m^{(i)}(\boldsymbol{\alpha};\boldsymbol{\alpha}^0).$$

We construct the predictor $V'_m(\boldsymbol{\alpha};\boldsymbol{\alpha}^0)$ such that it satisfies Eq. (23), i.e.,

$$V'_m(\boldsymbol{\alpha};\boldsymbol{\alpha}^0)=V_m(\boldsymbol{\alpha}^0)+W_m(\boldsymbol{\alpha};\boldsymbol{\alpha}^0)-W_m(\boldsymbol{\alpha}^0;\boldsymbol{\alpha}^0) \tag{29}$$

where $V_m(\boldsymbol{\alpha}^0)$ are found by solving Eq. (22). We define our iterative approach to determining the parameters $\boldsymbol{\alpha}$, for which the coefficients $V_m(\boldsymbol{\alpha})$ are zero, as follows (see also Fig. 1):

(1) Given the set of parameters $\boldsymbol{\alpha}^n$, construct the wave function $\Psi(\boldsymbol{\alpha}^n)$.

(2) Evaluate the required expectation values in Eqs. (22) and (27), using $\Psi(\boldsymbol{\alpha}^n)$, to find the fitting coefficients $V_m(\boldsymbol{\alpha}^n)$ and numerical functions $v_m^{(i)}(\boldsymbol{\alpha};\boldsymbol{\alpha}^n)$.

(3) *If* the total energy fitting coefficients are zero, i.e., $V_m(\boldsymbol{\alpha}^n)=0$ for all m, *then* we are done, *otherwise* continue.

(4) Construct the predictor function $V'_m(\boldsymbol{\alpha};\boldsymbol{\alpha}^n)$ in Eq. (29) using the fitted terms from step (2), and find the solution $\boldsymbol{\alpha}^{n+1}$ to the system,

$$V'_m(\boldsymbol{\alpha}^{n+1};\boldsymbol{\alpha}^n)=0 \quad \text{for all } m, \tag{30}$$

using the Newton-Raphson method,[27] (see Appendix C). Use this set of parameters $\boldsymbol{\alpha}^{n+1}$ in step (1).

Iterations continue until the total energy coefficients $V_m(\boldsymbol{\alpha}^n)$ tend to zero, and the values of the parameters $\boldsymbol{\alpha}^n$ converge. Note that even though the predictor $V'_m(\boldsymbol{\alpha};\boldsymbol{\alpha}^n)$ is only an approximation to the exact function $V_m(\boldsymbol{\alpha})$, Eq. (23) guarantees that at the converged solution

$$0=V'_m(\boldsymbol{\alpha};\boldsymbol{\alpha})=V_m(\boldsymbol{\alpha}).$$

In other words, the parameter set $\boldsymbol{\alpha}$ solves the Euler-Lagrange equations for $\langle\mathcal{H}\rangle$ exactly. Clearly, the larger the neighborhood within which the approximate relation in Eq. (24) holds, the faster this iterative procedure will converge. In the trivial case, if the exact analytic form of $V_m(\boldsymbol{\alpha})$ were known *a priori*, then we could just solve the Euler-Lagrange equations in one step using a suitable root-finding method.

Rather than starting the procedure from an initial guess $\boldsymbol{\alpha}=\mathbf{0}$, we may begin at step (4) using only the analytic terms in the predictor, since these are independent of the wave function and so do not require fitting. That is, we solve the system

$$S_m(\boldsymbol{\alpha}^1) = 0 \quad \text{for all } m.$$

The solution set $\boldsymbol{\alpha}^1$ is then used in step (1).

D. Partitioning the local energy

We note that the potential energy operators V_{ext} and V in the Hamiltonian $\mathcal{H}$, defined in Eq. (17), are multiplicative, and therefore their contributions to $E(\mathbf{R})$ are constant with respect to variations of the wave function parameters $\boldsymbol{\alpha}$. Variations of $\boldsymbol{\alpha}$ affect only the contributions of the differential kinetic energy operator. Thus solving the Euler-Lagrange equations for $\langle\mathcal{H}\rangle$ [Eq. (16)] amounts to adjusting the systematic fluctuations of the kinetic energy to cancel those in the potential energy exactly.

If we extract a variational part $\phi(\boldsymbol{\alpha})$ of Ψ, such that $\Psi = \phi\Psi'$, where ϕ is dependent on the set of parameters $\boldsymbol{\alpha}$, and Ψ' is independent of them. Then we may partition $E(\mathbf{R})$ as

$$E(\mathbf{R}) = \epsilon^{(1)}(\mathbf{R}) + \epsilon^{(2)}(\mathbf{R}) + \epsilon^{(3)}(\mathbf{R}),$$

where we define

$$\epsilon^{(1)}(\mathbf{R}) \equiv -\frac{1}{2}\sum_{i=1}^{N}\frac{\nabla_i^2\phi(\mathbf{R})}{\phi(\mathbf{R})},$$

$$\epsilon^{(2)}(\mathbf{R}) \equiv -\sum_{i=1}^{N}\frac{\boldsymbol{\nabla}_i\phi(\mathbf{R})}{\phi(\mathbf{R})}\cdot\frac{\boldsymbol{\nabla}_i\Psi'(\mathbf{R})}{\Psi'(\mathbf{R})},$$

$$\epsilon^{(3)}(\mathbf{R}) \equiv -\frac{1}{2}\sum_{i=1}^{N}\frac{\nabla_i^2\Psi'(\mathbf{R})}{\Psi'(\mathbf{R})} + \sum_{i=1}^{N} V_{\text{ext}}(\mathbf{r}_i) + \sum_{i<j}^{N} V(r_{ij}).$$

Clearly, $\epsilon^{(3)}$ is constant with respect to variations of $\boldsymbol{\alpha}$. Further analysis of each of $\epsilon^{(1)}$ and $\epsilon^{(2)}$ is necessary to determine how each depends on $\boldsymbol{\alpha}$. However, for ϕ expressible in the form of the Jastrow factors in Sec. II, i.e.,

$$\phi(\boldsymbol{\alpha}) = \exp\left[\sum_l \alpha_l \mathcal{O}_l\right],$$

where $\mathcal{O}_l$ are defined in Eq. (19), we see that (i) $\epsilon^{(2)}$ is at most linear in $\boldsymbol{\alpha}$, but involves terms coming from Ψ', so that it may be impossible to determine analytic expressions for the coefficients $v_m^{(2)}(\boldsymbol{\alpha})$; and (ii) $\epsilon^{(1)}$ is at most quadratic in $\boldsymbol{\alpha}$, and involves only ϕ, for which we have an analytic expression, and therefore may derive analytic approximations to the coefficients $v_m^{(1)}(\boldsymbol{\alpha})$ (see Sec. III E).

E. Analytic terms in the predictor

The initial predictor $S_m(\boldsymbol{\alpha})$ used in the iterative method involves only analytic terms. We determine these by direct expansion of particular local energy contributions, given the analytic form of the wave function. We now consider the contributions of the one- and two-body terms individually.

1. One-body Jastrow factor

Replacing $\phi(\boldsymbol{\alpha})$ in Sec. III D with $J_{1\text{b}}(\boldsymbol{\chi})$, we derive an analytic expression for the energy contribution $\epsilon^{(1)}$ (see Appendix D). From this we may extract those coefficients $v_{\mathbf{G}}^{(1)}$ of the functions $\mathcal{O}_{\mathbf{G}} = \Delta\rho_{\mathbf{G}}^{*}$:

$$v_{\mathbf{G}}^{(1)}(\boldsymbol{\chi}) = \frac{1}{2}G^2\chi_{\mathbf{G}} + \frac{1}{2}\sum_{\mathbf{G}'}\chi_{\mathbf{G}-\mathbf{G}'}(\mathbf{G}-\mathbf{G}')\cdot\mathbf{G}'\chi_{\mathbf{G}'}.$$

By assumption, the one-body contribution of $\epsilon^{(3)}$ is zero. That is, the mean-field methods used to calculate the determinant D should remove (approximately) all systematic one-body fluctuations in the local energy, and the use of "fluctuation functions" $\Delta\rho_{\mathbf{q}}$ in the two-body Jastrow factor approximately removes its effect on one-body operators. Therefore, we assume initially that the coefficients $v_{\mathbf{G}}^{(3)} = 0$. We are unable to derive an analytic expression for the coefficients $v_{\mathbf{G}}^{(2)}$ of the energy contribution $\epsilon^{(2)}$ and so, for the moment, we leave these aside. Constructing the full analytic approximation

$$S_{\mathbf{G}}(\boldsymbol{\chi}) = v_{\mathbf{G}}^{(1)}(\boldsymbol{\chi}),$$

we see that the roots of $S_{\mathbf{G}}$ are trivially $\boldsymbol{\chi} = \mathbf{0}$, as we would expect.

2. Two-body Jastrow factor

According to Sec. II, the two-body Jastrow correlation factor is divided into a short-range term J_{sr} and an inhomogeneous term J_{ih}. We optimize the variational parameters $\mathbf{u}$ in J_{ih}. As for the one-body Jastrow, we expand the corresponding energy contribution $\epsilon^{(1)}$ which depends on J_{ih} alone (see Appendix E). This leads us to an approximation to the coefficients of the functions $\mathcal{O}_{\mathbf{qq}'} = -P_{\mathbf{qq}'}$ in $\epsilon^{(1)}$:

$$v_{\mathbf{qq}'}^{(1)}(\mathbf{u}) = \frac{1}{2}u_{\mathbf{qq}'}(q^2+q'^2) + 2\sum_{\mathbf{kk}'} u_{\mathbf{qk}}(\mathbf{k}\cdot\mathbf{k}')\langle\rho_{\mathbf{k}'-\mathbf{k}}\rangle u_{\mathbf{k}'\mathbf{q}'}.$$

Again, we are unable to derive expressions for the coefficients corresponding to $\epsilon^{(2)}$. We extract a two-body contribution from the constant contribution $\epsilon^{(3)}$ (see Appendix E), as $v_{\mathbf{qq}'}^{(3)} \approx -\frac{1}{2}V(\mathbf{q})\delta(\mathbf{q}'-\mathbf{q})$ for the electron interaction V. We replace the true interaction in this expression with a pseudointeraction V_{ps}, which is generated for a given cutoff radius r_c and reference eigenvalue ϵ, as explained in Ref. 24. V_{ps} is used to generate the short-range two-body function u_{sr} [Eq. (4)], used in J_{sr}. The purpose of this modification of $v_{\mathbf{qq}'}^{(3)}$ is to account for the presence of J_{sr} in Ψ and is explained in Appendix F.

From these contributions we construct the analytic predictor

$$S_{\mathbf{qq}'}(\mathbf{u}) = v_{\mathbf{qq}'}^{(1)}(\mathbf{u}) - \frac{1}{2}V_{\text{ps}}(\mathbf{q})\delta(\mathbf{q}-\mathbf{q}').$$

We notice that for periodic systems, this function is separable in the points $\mathbf{q}$ of the first Brillouin zone (BZ). For each $\mathbf{q}$ in the BZ, we may expand $S_{\mathbf{q}+\mathbf{G},\mathbf{q}+\mathbf{G}'}$ as a function of $\{u_{\mathbf{q}+\mathbf{G},\mathbf{q}+\mathbf{G}'}\}$, with no coupling to parameters $u_{\mathbf{q}'+\mathbf{H},\mathbf{q}'+\mathbf{H}'}$

for $\mathbf{q}'\neq\mathbf{q}$. This block diagonal form of the analytic predictor allows us to find the roots $\mathbf{u}^1$ by solving for each block (i.e., each $\mathbf{q}$) individually. We do this using the Newton-Raphson method.

A reliable initial guess for the Newton-Raphson method, rather than using $\mathbf{u}=\mathbf{0}$, is the homogeneous solution of $S_{\mathbf{qq}'}=0$. If we regard the system as homogeneous, i.e., $\langle\rho_{\mathbf{G}}\rangle=0$ for $\mathbf{G}\neq\mathbf{0}$, then the solution is

$$u_{\mathbf{qq}'}=\delta(\mathbf{q}'-\mathbf{q})\,\frac{1}{4N}\left[\sqrt{1+\frac{8NV_{\mathrm{ps}}(\mathbf{q})}{q^2}}-1\right]. \qquad (31)$$

This $u_{\mathbf{qq}'}$ is used as the starting point for the Newton-Raphson iterations, and for all systems studied, this initial guess produced convergent roots of $S_{\mathbf{qq}'}=0$.

F. Numerical terms in the predictor

To complete the construction of the predictor $V'_m(\boldsymbol{\alpha};\boldsymbol{\alpha}^0)$ defined in Eq. (29), for a given set of variational parameters $\boldsymbol{\alpha}^0$, we must determine some terms numerically by fitting fluctuations in the energy of the system. We calculate $V_m(\boldsymbol{\alpha}^0)$ in Eq. (29) by fitting the entire local energy, using Eq. (22). Also, we use the fitting method to numerically determine the functions $v_m^{(2)}(\boldsymbol{\alpha};\boldsymbol{\alpha}^0)$ given in Eq. (28), by fitting the energy contribution $\boldsymbol{\epsilon}^{(2)}$ and each of its derivatives with respect to the parameters to be optimized. Again, we note that $\boldsymbol{\epsilon}^{(2)}$ is explicitly linear in the parameters $\boldsymbol{\alpha}$ and the approximation in Eq. (28) is exact in this case.

To do all this numerical work we use Monte Carlo sampling to determine the required expectation values

$$\langle\Delta\mathcal{O}_m\Delta\mathcal{O}_n\rangle,$$

$$\langle\Delta E\Delta\mathcal{O}_m\rangle,$$

$$\langle\Delta\boldsymbol{\epsilon}^{(2)}\Delta\mathcal{O}_m\rangle,$$

$$\left\langle\Delta\left(\frac{\partial\boldsymbol{\epsilon}^{(2)}}{\partial\alpha_m}\right)\Delta\mathcal{O}_n\right\rangle,$$

for m and n ranging over the number of parameters N_α in the set $\boldsymbol{\alpha}$. However, for a simultaneous optimization of one- and two-body terms, we make some simplifications to reduce the computational workload. If we assume that the parameter α_m can be varied independently of α_n, this corresponds to assuming that $\langle\Delta\mathcal{O}_m\Delta\mathcal{O}_n\rangle=0$.

In the expansion of the analytic local energy terms coming from J_{ih}, outlined in Appendix E, we saw that these consisted of one- and two-body terms. However, the one-body terms contained Fourier coefficients of the average charge density $\langle\rho_{\mathbf{q}}\rangle$, which are zero for $\mathbf{q}\neq\mathbf{G}$, a reciprocal lattice vector. Also, the analytic form of the predictor (Sec. III E) is separable in the k points of the first Brillouin zone. Therefore, we make the following approximations to the covariance matrix $\langle\Delta\mathcal{O}_m\Delta\mathcal{O}_n\rangle$,

$$\langle\Delta\mathcal{O}_{\mathbf{q}+\mathbf{G},\mathbf{q}+\mathbf{G}'}\Delta\mathcal{O}_{\mathbf{q}'+\mathbf{H}',\mathbf{q}'+\mathbf{H}'}\rangle=0 \quad \text{for } \mathbf{q}\neq\mathbf{q}',$$

$$\langle\Delta\mathcal{O}_{\mathbf{q}+\mathbf{G},\mathbf{q}+\mathbf{G}'}\Delta\mathcal{O}_{\mathbf{H}}\rangle=0 \quad \text{for } \mathbf{q}\neq 0,$$

for reciprocal lattice vectors $\mathbf{G},\mathbf{G}',\mathbf{H}$, and $\mathbf{H}'$.

In other words, we regard one- and two-body optimizations as independent for nonzero k points in the first Brillouin zone.[46] If we also exclude the covariance terms between one- and two-body operators for $\mathbf{q}=\mathbf{0}$, we find that this slows the convergence of the method for smaller systems. For larger systems, this exclusion prevents the convergence of the Newton-Raphson method at the first iteration of our method, thus halting the optimization process. However, we are not, in any way, confined to using the Newton-Raphson method to find the roots of the predictor, and other, more robust, root-finding methods might overcome this problem.

Note that $\boldsymbol{\epsilon}^{(2)}$ depends on the variational part ϕ of the wave function which is being optimized. The variational components are $J_{1\mathrm{b}}$ and $J_{\mathrm{ih}}(\mathbf{q})$ for each $\mathbf{q}$ in the BZ, where we define

$$J_{\mathrm{ih}}(\mathbf{q})=\exp\left[-\sum_{\mathbf{GG}'}u_{\mathbf{q}+\mathbf{G},\mathbf{q}+\mathbf{G}'}P_{\mathbf{q}+\mathbf{G},\mathbf{q}+\mathbf{G}'}\right],$$

so that $J_{\mathrm{ih}}=\Pi_{\mathbf{q}}J_{\mathrm{ih}}(\mathbf{q})$. This greatly reduces the complexity of the predictor, without sacrificing convergence of the method for the systems studied here. Ultimately, the predictor $V'_m(\boldsymbol{\alpha};\boldsymbol{\alpha}^0)$ is itself only an approximation of the true function $V_m(\boldsymbol{\alpha})$, but by definition matches this function at the current values of the parameters being optimized. Therefore, approximations in the predictor affect only the rate of convergence of the method.

The expressions for $\boldsymbol{\epsilon}^{(2)}$ and its derivatives $\partial\boldsymbol{\epsilon}^{(2)}/\partial\chi_{\mathbf{G}}$ for the one-body Jastrow factor $J_{1\mathrm{b}}$ are given in Appendix D. Upon fitting these terms to the operators $\mathcal{O}_{\mathbf{G}}=\Delta\rho^*_{\mathbf{G}}$, we may construct the function $v_m^{(2)}(\boldsymbol{\alpha};\boldsymbol{\alpha}^0)$. This defines the fitted terms $T_{\mathbf{G}}$ in the predictor (Sec. III C).

For a given $\mathbf{q}$ in the BZ, we use the expressions for $\boldsymbol{\epsilon}^{(2)}$ and the derivatives $\partial\boldsymbol{\epsilon}^{(2)}/\partial u_{\mathbf{q}+\mathbf{G},\mathbf{q}+\mathbf{G}'}$, given in Appendix E, corresponding to the two-body Jastrow factor $J_{\mathrm{ih}}(\mathbf{q})$ defined above. We construct the function $v^{(2)}_{\mathbf{q}+\mathbf{G},\mathbf{q}+\mathbf{G}'}(\mathbf{u}_{\mathbf{q}})$, where $\mathbf{u}_{\mathbf{q}}=\{u_{\mathbf{q}+\mathbf{G},\mathbf{q}+\mathbf{G}'};\ \text{for } \mathbf{G},\mathbf{G}'\}$, and define the numerical predictor terms $T_{\mathbf{q}+\mathbf{G},\mathbf{q}+\mathbf{G}'}=v^{(2)}_{\mathbf{q}+\mathbf{G},\mathbf{q}+\mathbf{G}'}$, which contribute to the linear dependence of the predictor on $\mathbf{u}_{\mathbf{q}}$.

IV. RESULTS

We now apply this optimization method to diamond and rhombohedral graphite. We shall compare the correlation factors determined in both these systems, given the inhomogeneity and anisotropy of the electron charge density in graphite, relative to diamond.

The convergence criterion of our iterative optimization: $V_m(\boldsymbol{\alpha}^n)=0$, requires the examination of possibly thousands of parameters, and it is difficult to visualize the overall convergence of the method. For illustrative purposes, we use the coefficients $V_m(\boldsymbol{\alpha}^n)$ to construct a single function V^n. In the chosen Fourier basis, this amounts to reconstructing the real-space function V^n from its Fourier coefficients. For one-body optimizations

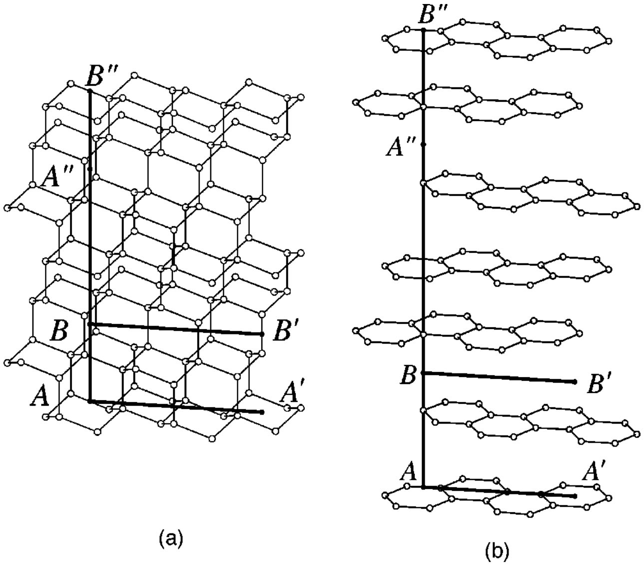

FIG. 2. Crystal structure of (a) diamond and (b) rhombohedral graphite, illustrating stacked layers of hexagonally arranged carbon atoms for graphite and buckled layers for diamond. In both structures, the point A lies at the midpoint of a carbon-carbon bond, with the lines AA' and AA'' extending within a layer and perpendicular to the layers, respectively. The point B lies midway between two layers (at a hexagonal interstitial point in diamond) and the lines BB' and BB'' extend between the layers and perpendicular to the layers, respectively.

$$V^{n}(\mathbf{r})=\sum_{\mathbf{G}} V_{\mathbf{G}}(\boldsymbol{\chi}^{n})e^{i\mathbf{G}\cdot\mathbf{r}}, \tag{32}$$

and for two-body optimizations

$$V^{n}(\mathbf{r},\mathbf{r}')=\sum_{\mathbf{q}}\sum_{\mathbf{GG}'} e^{-i(\mathbf{q}+\mathbf{G})\cdot\mathbf{r}}V_{\mathbf{q}+\mathbf{G},\mathbf{q}+\mathbf{G}'}(\mathbf{u}^{n})e^{i(\mathbf{q}+\mathbf{G}')\cdot\mathbf{r}'}. \tag{33}$$

In practice, we monitor the decrease of these functions, but they will never converge to zero due to inescapable statistical fluctuations that result from the finite sampling used in the VMC approach to the least-squares fitting (Sec. V C and Appendix G).

Two-body functions, such as $V^{n}(\mathbf{r},\mathbf{r}')$ and the Jastrow correlation function $u(\mathbf{r},\mathbf{r}')$, are functions of six variables and so, extracting useful information from them is difficult. For illustrative purposes, we indicate in Fig. 2 two points A and B in both the diamond and rhombohedral graphite structures, corresponding to high and low electron charge density regions, respectively. A lies midway between two bonded carbon atoms and B lies midway between two layers of carbon atoms. We shall position the first electron at either A or B. The second electron shall be moved away from this position along one of the following line segments (indicated by heavy black lines in Fig. 2): AA' lying within a layer of carbon atoms; AA'' perpendicular to the layers; BB' lying between two layers; and BB'' perpendicular to the layers.

By this means we may plot inhomogeneous two-body functions in terms of the relative separation of the electrons in the system. In particular, we may draw some conclusions about electron correlation in the system by an examination of u. We may determine the isotropy of u by comparing plots of u with the first electron kept at the same point but the second electron moved in perpendicular directions, e.g., by comparing plots designated by AA' and AA''. The homogeneity of u may be seen by comparing plots of u with the second electron moving in parallel directions from different positions of the first electron, e.g., by comparing AA' and BB'. Any differences between these plots of u are attributable to the inhomogeneity and anisotropy of the electron correlation factors in the systems studied.

We use periodic boundary conditions (PBC's) to approximate the infinite crystal.[4] The simulation cells consist of an $N_1\times N_2\times N_3$ unit cell arrangement. The unit cells in each system are defined by the Bravais lattice basis vectors. For diamond, we use the basis $\{(a/2,a/2,0);(0,a/2,a/2);(a/2,0,a/2)\}$, where $a=6.72$ a.u., corresponding to a carbon bond length of 2.91 a.u. For rhombohedral graphite we use the basis $\{(0,a,c);(-\sqrt{3}a/2,-a/2,c);(\sqrt{3}a/2,-a/2,c)\}$, where $a=2.68$ a.u. is the bond length within the layers, and $c=6.33$ a.u. is the layer separation. These are experimentally determined structural parameters.[28]

We construct the Slater determinant D for both systems using DFT calculations in the LDA.[25] The LDA orbitals are expanded using a linear combination of atomic orbitals (LCAO) centered on each of the two carbon atoms in the unit cell. These LDA/LCAO calculations follow the scheme of Chan *et al.*[29] where, initially, the orbitals and ionic potential are expanded in a basis of localized functions (Gaussians defined by a set of decay constants) and then self-consistent corrections to the Hartree and exchange-correlation potentials are determined in Fourier space. This scheme uses a much smaller basis of planewaves than methods which require the expansion of the orbitals and ionic potential entirely in planewaves.

A. Removal of cusp

In Sec. II A, we discussed the advantages of removing the short-range cusp from the function u in order to improve its representability as a linear combination of smooth functions. Figure 3(a) compares the pseudointeraction V_{ps}, generated using a cutoff radius of $r_c=1.9$ a.u. and energy eigenvalue $\epsilon=0.2$ hartree (as discussed in Ref. 24), with the Coulomb interaction $V=e^2/r$. (In atomic units $e^2=1$). The pseudointeraction is used to generate a short-range Jastrow funtion u_{sr}, which is shown in Fig. 3(b) for the relative angular momenta $l=1$ and $l=0$ corresponding to parallel spin and anti-parallel spin correlations, respectively. The short-range Jastrow factor used in all subsequent calculations is that generated with these particular values of r_c and ϵ. Subsequent figures in this paper, which involve u_{sr}, represent antiparallel spin correlation only.

The cutoff required for a convergent Fourier expansion of a smooth cuspless function should be much less than that required for a function with a short-range cusp. Therefore, using a cuspless form greatly reduces the number of terms required to represent the inhomogeneous form of the Jastrow factor accurately in Fourier space. We illustrate this point

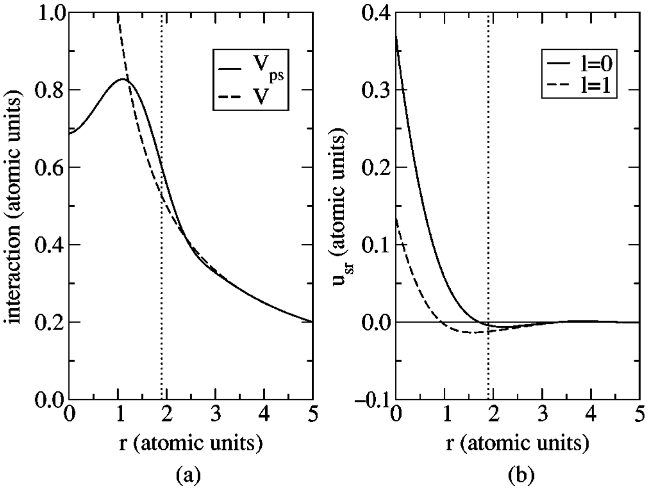

FIG. 3. (a) The pseudointeraction V_{ps} (solid line) and the Coulomb interaction $V=1/r$ (dashed line) vs electron separation r. (b) The short-range Jastrow function u_{sr}, generated from V_{ps}, vs electron separation r, for angular momenta $l=0$ (solid line) and $l=1$ (dashed line). V_{ps} is generated using $r_c=1.9$ a.u. (indicated by vertical dotted line) and $\epsilon=0.2$ hartree (see the text).

using a simple example. In Fig. 4 we plot q^2 times the Fourier transform, for wave vector q, of the Yukawa-style homogeneous correlation factor

$$u_{\mathrm{h}}=\frac{A}{r}(1-e^{-r/F}), \tag{34}$$

which has been used by many authors to approximate electron correlation in a variety of systems.[4,10,19–21,30] We set $A=1$ a.u. (F is determined from A to satisfy the cusp conditions[22] and depends on the relative spin of the electrons). Also shown in Fig. 4 is q^2 times the Fourier transform of the cuspless difference $u_{\mathrm{h}}-u_{\mathrm{sr}}$. We assume that, for large electron separations, the electron correlation is approximately spin independent. Therefore, for each q, we plot the mean of the parallel spin and antiparallel spin values of the functions.

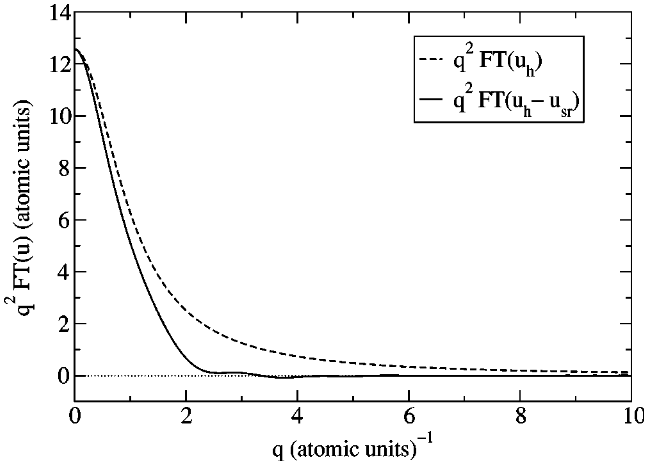

FIG. 4. Fourier transform times q^2 of (a) the homogeneous Jastrow factor u_{h} (dashed line) and (b) the cuspless difference $u_{\mathrm{h}}-u_{\mathrm{sr}}$ (solid line) vs wave vector q. The homogeneous Jastrow factor parameter $A=1$ a.u. [Eq. (34)].

In practice, we use the first zero of the Fourier transform of V_{ps} as the Fourier space cutoff k_c used in the definition of J_{ih} in Eq. (14). For $r_c=1.9$ a.u. we use the cutoff $k_c=2.185$ (a.u.)$^{-1}$, beyond which the Fourier transform of the cuspless u function is approximately zero (Fig. 4). Combining both the short-range and inhomogeneous forms of the Jastrow factor using this scheme produces a form that is approximately independent of the cutoff, since decreasing r_c increases the reciprocal space cutoff k_c.

B. Homogeneous Yukawa-style Jastrow factor

For comparison with the inhomogeneous u functions determined in the following sections we use the Yukawa-style homogeneous function u_{h} of Eq. (34) and construct a homogeneous trial wave function of the form $\Psi=J_{\mathrm{sr}}J_{\mathrm{h}}J_{1\mathrm{b}}D$. As with the inhomogeneous trial wave function, we represent short-range correlations (i.e., the cusp) using J_{sr}, and represent the remaining correlations using a homogeneous Jastrow factor

$$J_{\mathrm{h}}=\exp\left[-\sum_{q<k_c}\tilde{u}(\mathbf{q})P_{\mathbf{qq}}\right],$$

where we define

$$\tilde{u}(\mathbf{q})\equiv\int_{\Omega}[u_{\mathrm{h}}(r)-u_{\mathrm{sr}}(r)]e^{-i\mathbf{q}\cdot\mathbf{r}}d\mathbf{r}.$$

For the uniform electron gas, Bohm and Pines[18] predicted that the true function u should decay as $1/\omega_p r$, at large separations r, where ω_p is the plasma frequency. Rather than use this limiting value $A=1/\omega_p$ in Eq. (34), it is common to treat A as a free parameter such that the energy is minimized. Using variational calculations we can determine the optimal value of A.[21] We optimize the one-body Jastrow factor $J_{1\mathrm{b}}$ in Eq. (15), using our iterative method.

In Fig. 5 we plot u_{h} and compare it with the reconstructed function

$$u(\mathbf{r})=u_{\mathrm{sr}}(r)+\sum_{q<k_c}\tilde{u}(\mathbf{q})e^{i\mathbf{q}\cdot\mathbf{r}} \tag{35}$$

for $3\times3\times3$ unit cell simulations of diamond and rhombohedral graphite. Periodic boundary conditions and the anisotropy of the unit cell make this reconstructed form appear quite different from the original isotropic function u_{h}.

In particular, for rhombohedral graphite the unit cell used is quite anisotropic, leading to marked differences in the reconstructed function along the perpendicular line segments AA' and AA''. (The homogeneity of u_{h} is preserved and so we only plot the function for point A, since all other points are equivalent.) Note that the Jastrow u function is defined up to an arbitrary constant, much like a potential, since this constant affects only the wave function normalization and contributes nothing to the description of correlation. Therefore, it is of no consequence that the functional form of u_{h} in

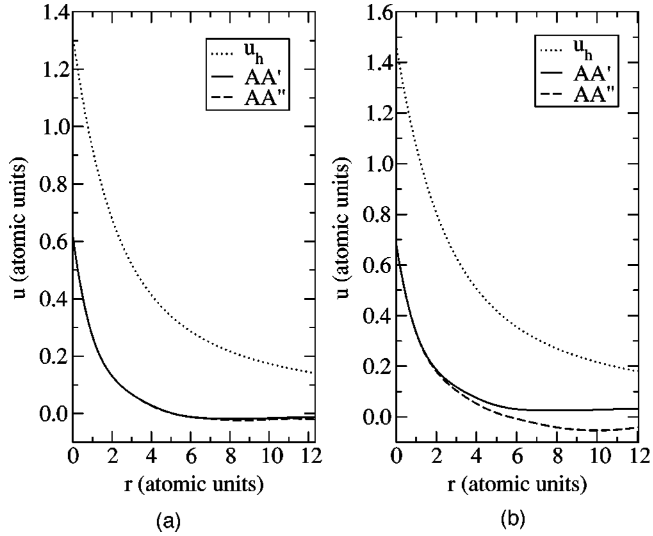

FIG. 5. Jastrow correlation factors vs electron separation r for (a) diamond and (b) rhombohedral graphite. (i) The function u_{h} (dotted line) as defined in Eq. (34) with optimized parameter $A = 1.739$ for diamond and 2.170 for graphite. (ii) The reconstructed function u of Eq. (35), from a $3\times3\times3$ simulation region in both systems, for electron separations along the line segments AA' (solid line) and AA'' (dashed line) from Fig. 2.

Eq. (34) appears shifted above the reconstructed forms compatible with PBC's. This is due to the removal of the constant Fourier coefficient of the correlation function $u(\mathbf{G}=0)$ from the expansion of the reconstructed function in Eq. (35). We note that the cusp conditions[22] are maintained by all forms.

We optimize the one-body Jastrow $J_{1\mathrm{b}}$ using the method described in Sec. III. For the diamond simulations we used a Fourier space cutoff $G_c = 5.0$ (a.u.)$^{-1}$, giving 84 variational one-body parameters. This is more than enough for an accurate representation of one-body terms in the wave function (see Sec. V A). For graphite simulations we used $G_c = 3.1$ (a.u.)$^{-1}$, giving 30 variational one-body parameters.

The values of the electronic energy per atom for various optimizations of this homogeneous trial wave function are shown in Figs. 6 and 7. Each iteration involves averaging over 10^5 Monte Carlo samples. However, this amount of averaging is more than enough for an accurate implementation of the method. The optimization of the one-body terms in the $3\times3\times3$ simulation of graphite (Fig. 7) involves only 2.5×10^4 samples per iteration, and is well converged.

Specific values of the total energy and its associated variance during the optimization process, for the largest simulations of diamond and graphite are presented in Table II. On average, the gain in energy following one-body optimization is approximately 1.0 mhartree/atom for diamond and 2.5 mhartree/atom for graphite. We note that the necessity for a one-body correction is a consequence of the inhomogeneity of the electronic charge density in the system,[20,21] and it is not surprising that the gain in energy is larger for the more inhomogeneous system, graphite, than for diamond. We notice that there is no significant decrease in the variance for this one-body optimization. For the smaller systems, the decrease is approximately 2% and is negligible for the larger systems.

The Slater determinant used in the diamond calculations of Fig. 6 is composed of single-particle orbitals generated from the LDA/LCAO scheme already described.[29] The localized functions use a Gaussian basis with three decays of 0.24, 0.797, and 2.65, and the orbitals possess s, p, and d symmetries. The exchange-correlation functional used was of the Ceperley-Alder[8] form. The cutoff for the Fourier space expansion of the charge density and the iterative corrections to the Hartree and exchange-correlation potential during the self-consistent LDA calculation was 64 Rydberg. Evaluation of the orbitals of the Slater determinant within the VMC sampling was achieved using a mixed-basis approach[4] with a real-space cutoff of 3.5 a.u. and a reciprocal-space cutoff of 5.7 (a.u.)$^{-1}$.

The graphite calculations shown in Fig. 7 use a Slater determinant generated using a localized basis set with s and p symmetries only. Four Gaussian decays were used: 0.19, 0.474, 1.183, and 2.95. The orbitals were generated from

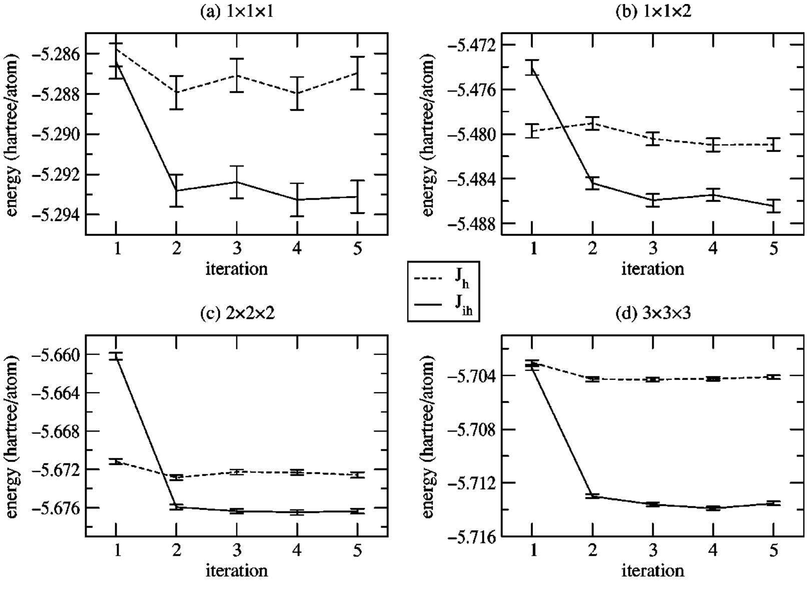

FIG. 6. Energy of diamond, in Hartree/atom, for each iteration of the optimization process and various simulation cell sizes: (a) $1\times1\times1$ unit cell, with $N=8$ electrons; (b) $1\times1\times2$, $N=16$; (c) $2\times2\times2$, $N=64$; (d) $3\times3\times3$, $N=216$. Optimizations using different correlation factors are shown: $J = J_{\mathrm{sr}}J_{\mathrm{h}}J_{1\mathrm{b}}$, Yukawa-style homogeneous form (dashed line) and $J = J_{\mathrm{sr}}J_{\mathrm{ih}}J_{1\mathrm{b}}$ inhomogeneous form (solid line). Total energy estimates for each iteration are calculated using 10^5 Monte Carlo samples. Error bars are equal to the standard error of the mean value of the energy.

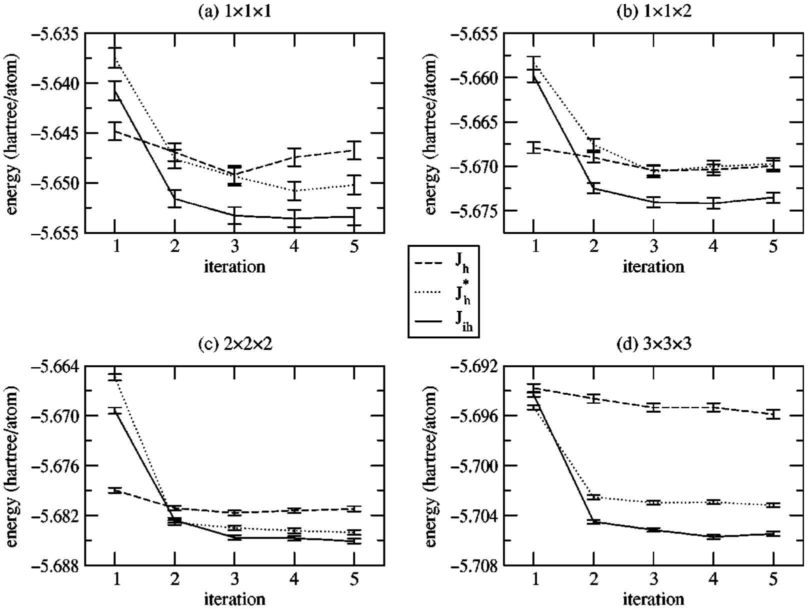

FIG. 7. Energy of rhombohedral graphite, in Hartree/atom, for each iteration of the optimization process and various simulation cell sizes: (a) $1\times1\times1$ unit cell, with $N=8$ electrons; (b) $1\times1\times2$, $N=16$; (c) $2\times2\times2$, $N=64$; (d) $3\times3\times3$, $N=216$. Optimizations using different correlation factors are shown: $J=J_{sr}J_hJ_{1b}$, Yukawa-style homogeneous form (dashed line); $J=J_{sr}J_h^*J_{1b}$, optimal homogeneous form (dotted line); and $J=J_{sr}J_{ih}J_{1b}$ inhomogeneous form (solid line). Total energy estimates for each iteration are calculated using 10^5 Monte Carlo samples. (The optimization of the Yukawa-style homogeneous form for the $3\times3\times3$ simulation used only 2.5×10^4 samples per iteration.)

LDA calculations incorporating the Hedin-Lundqvist[31] exchange-correlation functional, and a cutoff of 36 Rydberg for the Fourier space expansion of the charge density and correction to the potential. The same mixed-basis cutoffs used in the diamond calculations were used here, for evaluation of the orbitals in the Slater determinant during VMC sampling. (See Sec. V A for a discussion of basis-set convergence of the total energy in graphite.)

Figure 8 illustrates the convergence during optimization of the one-body function $\chi(\mathbf{r})$ in Eq. (1) (where χ is the accumulation of all one-body terms from each of the Jastrow factors) for the $3\times3\times3$ diamond calculation. The optimized function is statistically well-determined, and the change from the initial one-body function, $\chi^1=\chi^0$ of Eq. (8) (as defined in Refs. 20 and 21), is well defined. This alteration of the one-body function may be compared to similar calculations performed using variance minimization.[20,21]

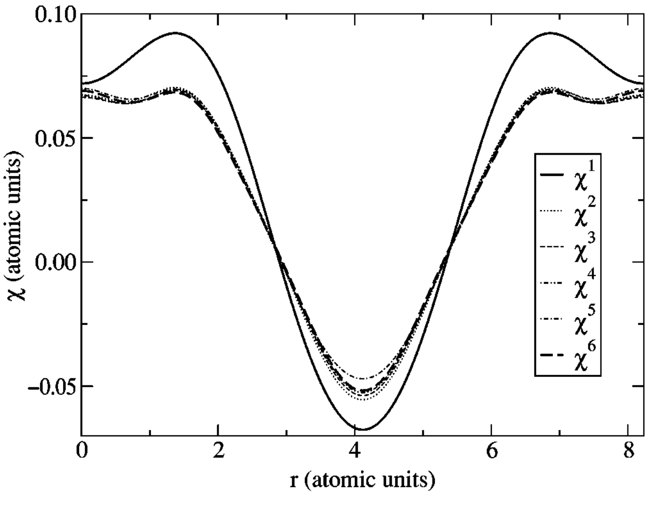

FIG. 8. The diamond one-body function χ vs position r along the line segment AA' [Fig. 2(a)]. χ^n indicates the one-body function used at iteration n of the optimization of J_{1b} in the presence of J_h in Fig. 6(d).

The number of parameters for optimization could be greatly reduced through exploitation of the crystal point-group symmetry of the structures involved. However, it is worth noting that the optimization process preserves the natural symmetry of the system (within statistical error) without such measures, illustrating that for nonsymmetric systems with large numbers of parameters, this optimization process should be quite robust.

The one-body function $V^n(\mathbf{r})$, reconstructed from the coefficients associated with the local energy $V_{\mathbf{G}}$ at iteration n according to Eq. (32), is shown in Fig. 9. Clearly, this function decreases in magnitude, indicating a decrease in the magnitudes of the Euler-Lagrange derivatives. Beyond the first iteration, $V^n(\mathbf{r})$ is of the same order of magnitude as its associated standard error, and so is statistically insignificant. Therefore, the method has essentially converged after only one iteration. The noisiest regions of $V^n(\mathbf{r})$ correspond to regions of low density where the Monte Carlo sampling is less frequent.

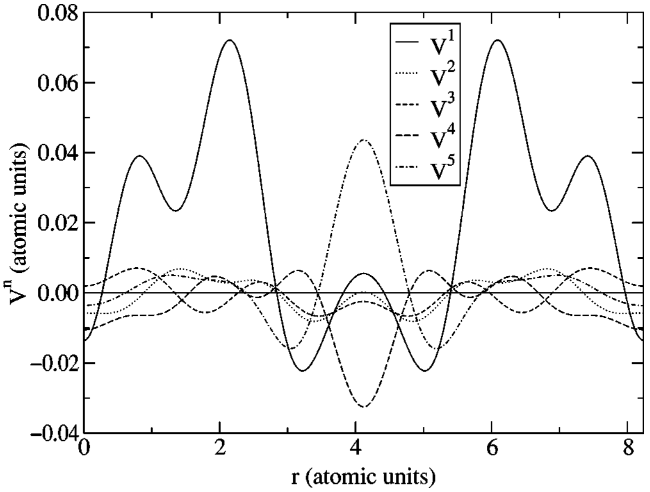

FIG. 9. The reconstructed one-body function V^n, defined in Eq. (32), vs position r as for Fig. 8.

TABLE I. The number of variational parameters in the correlation factors, used during the total energy calculations presented in Figs. 6 and 7, for diamond and rhombohedral graphite. Simulation cell sizes are described as $N_1 \times N_2 \times N_3$ unit cell arrangements. $J_{\rm 1b}$ is the one-body Jastrow factor; $J_{\rm h}^*$ is the optimal homogeneous correlation factor; $J_{\rm ih}$ is the optimal inhomogeneous correlation factor. (The Yukawa-style correlation factor $J_{\rm h}$ is not included as it contains only one variational parameter.)

simulation cell	$J_{\rm 1b}$	$J_{\rm h}^*$	$J_{\rm ih}$
		diamond	
1×1×1	84	−	56
1×1×2	84	−	112
2×2×2	84	−	448
3×3×3	84	−	1151
		graphite	
1×1×1	30	11	132
1×1×2	30	21	242
2×2×2	30	85	990
3×3×3	30	282	3106

TABLE II. The energies (E) and their variances (σ^2) determined at each step of the optimization process for the largest ($3 \times 3 \times 3$) simulations of diamond and rhombohedral graphite using various correlation factors. The standard error in the energy is computed from the variance as $\Delta E = \pm \sqrt{\sigma^2/N}$. The error in the variances (included in parentheses) is estimated assuming they have a chi-squared distribution. (All numerical values in Hartree atomic units per atom.)

J	Iteration	E	ΔE	σ^2
		diamond		
$J_{\rm sr}J_{\rm 1b}J_{\rm h}$	1	−5.703 030	0.000 155	0.002 41(2)
	2	−5.704 282	0.000 151	0.002 27(2)
	3	−5.704 306	0.000 151	0.002 29(2)
	4	−5.704 247	0.000 151	0.002 28(2)
	5	−5.704 118	0.000 150	0.002 26(2)
$J_{\rm sr}J_{\rm 1b}J_{\rm ih}$	1	−5.703 462	0.000 154	0.002 38(2)
	2	−5.713 010	0.000 139	0.001 92(2)
	3	−5.713 620	0.000 138	0.001 91(2)
	4	−5.713 910	0.000 139	0.001 92(2)
	5	−5.713 539	0.000 140	0.001 95(2)
		graphite		
$J_{\rm sr}J_{\rm 1b}J_{\rm h}$	1	−5.693 792	0.000 338	0.002 85(5)
	2	−5.694 640	0.000 336	0.002 83(5)
	3	−5.695 348	0.000 338	0.002 85(5)
	4	−5.695 337	0.000 337	0.002 85(5)
	5	−5.695 882	0.000 353	0.003 12(6)
$J_{\rm sr}J_{\rm 1b}J_{{\rm h}*}$	1	−5.695 342	0.000 166	0.002 75(2)
	2	−5.702 535	0.000 161	0.002 58(2)
	3	−5.702 982	0.000 161	0.002 59(2)
	4	−5.702 934	0.000 162	0.002 63(2)
	5	−5.703 191	0.000 163	0.002 67(2)
$J_{\rm sr}J_{\rm 1b}J_{\rm ih}$	1	−5.694 342	0.000 167	0.002 80(3)
	2	−5.704 501	0.000 158	0.002 49(2)
	3	−5.705 159	0.000 155	0.002 39(2)
	4	−5.705 730	0.000 156	0.002 44(2)
	5	−5.705 471	0.000 159	0.002 54(2)

C. Optimal homogeneous Jastrow factor

To compare directly with current homogeneous Jastrow factors,[32,33] we also include the optimization of a generalized homogeneous two-body function u. This u function is constructed according to Eq. (35), but we allow complete variational freedom in the terms $\tilde{u}(\mathbf{q})$. We optimize these $\tilde{u}(\mathbf{q})$ according to the method outlined in Sec. III, except that, of course, u is now homogenoeus (but still retains the variational freedom for anisotropy). Since this expansion is complete up to the cutoff k_c, it is equivalent to any other homogeneous representation of u and may be compared with current homogeneous correlation factors. However, where others have optimized the parameters in such correlation factors using variance minimization, our optimal homogeneous u minimizes the total energy of the system. We denote the optimal homogeneous part of the Jastrow factor as $J_{\rm h}^*$. The Jastrow factor $J = J_{\rm sr} J_{\rm h}^* J_{\rm 1b}$, combined with the Slater determinant D, is our best approximation of the true many-body eigenstate using a homogeneous two-body Jastrow factor.

Figure 7 illustrates the total energy of various graphite systems during the optimization of $J_{\rm h}^*$. The initial guess for $J_{\rm h}^*$, used in iteration 1, is a solution to the RPA equations for a homogeneous system. We note that the optimal homogeneous Jastrow factor produces total energies that are consistently lower than (or at least comparable to) those of the simpler Yukawa-style $J_{\rm h}$. This is an obvious consequence of the increased variational freedom of $J_{\rm h}^*$. The numbers of variational parameters $\tilde{u}(\mathbf{q})$ used for each system are displayed in Table I. The specific values of the total energy and its variance, during the optimization of the graphite $3 \times 3 \times 3$ system, are given in Table II.

D. Inhomogeneous RPA Jastrow factor

The analytic guess for the two-body predictor functions $S_{\mathbf{q}\mathbf{q}'}(\mathbf{u})$, outlined in Sec. III E 2, leads to an inhomogeneous generalization of the RPA equations. The solution to these equations is the function $u_1 = u_{\rm RPA}$, shown in Fig. 10 for the $3 \times 3 \times 3$ graphite simulation. We notice some inhomogeneity in $u_{\rm RPA}$ at intermediate- and long-range electron separations. A more homogeneous and isotropic $u_{\rm RPA}$ was found for diamond, as we would expect since diamond possesses a more uniform electron density than graphite.

In Figs. 6 and 7, the first point on all solid curves indicates the total energy per atom in each simulation calculated using $u_{\rm RPA}$. In comparison with the energy calculated using the homogeneous Jastrow function $u_{\rm h}$, we see that the inhomogeneous RPA trial wave function is at best comparable in accuracy with the optimized trial wave function with homogeneous two-body Jastrow factor, and often less accurate.

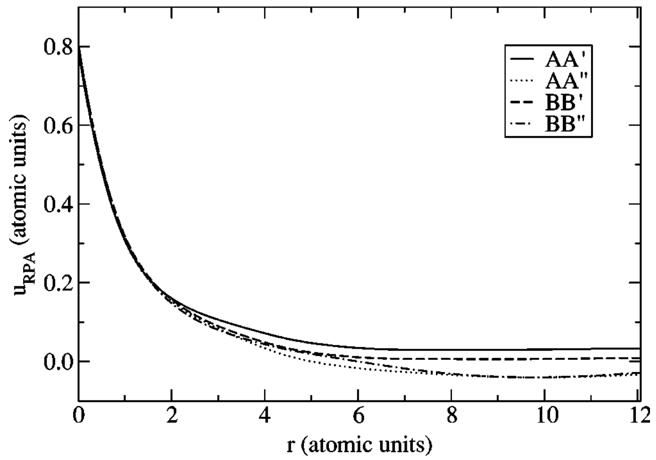

FIG. 10. The Jastrow factor u_{RPA} vs electron separation r along the line segments indicated in Fig. 2(b), for a $3\times3\times3$ simulation of rhombohedral graphite.

These results are different from those of Gaudoin *et al.*[20] for model systems: they find that their inhomogeneous generalization of the RPA produces wave functions that yield lower energies than the homogeneous form.

E. Optimal inhomogeneous Jastrow factor

We simultaneously optimized the parameters in both the one-body Jastrow factor $J_{1\mathrm{b}}$ and the fully inhomogeneous form of the two-body Jastrow factor J_{ih}, using our iterative method. The convergence of the total energy per atom for various simulations of diamond and rhombohedral graphite are shown in Figs. 6 and 7. (The Slater determinants used in combination with the homogeneous Jastrow factor J_{h} in Sec. IV B are also used here.) Convergence of the total energy is achieved in approximately three iterations in all cases and is stable. This is remarkable given that the system sizes range from eight electrons and 140 independent parameters to 216 electrons and 3136 independent parameters. Also, we used the same amount of Monte Carlo sampling, *viz.*, 10^5 samples per iteration, to determine the required expectation values for all simulations. In all cases the fully inhomogeneous form of u allows us to determine more accurate trial wave functions with substantially lower energies than the homogeneous trial functions. In general, the gain in energy through using an inhomogeneous rather than a homogeneous wave function is of the order of 5 mhartree/atom for both diamond and rhombohedral graphite. Specific values for the energy and its variance are given in Table II for the largest simulations.

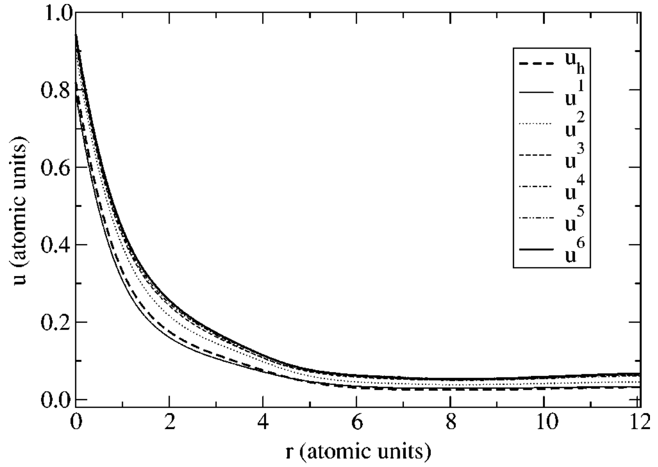

FIG. 11. Graphite Jastrow correlation functions u^n with respect to electron separation r on the line segment AA' (Fig. 2) for each iteration n during optimization of J_{ih} for a $3\times3\times3$ simulation region [Fig. 7(d)]. Also shown is the reconstruction of the homogeneous function u_h (heavy dotted line) defined in Eq. (35) with an optimized parameter $A=2.170$.

Figure 11 illustrates the rapid convergence of the two-body wave function parameters $\mathbf{u}$ to their optimal values during the largest graphite optimization ($3\times3\times3$) and is typical of all the optimizations performed in both diamond and graphite. Beyond the third iteration, no clear distinction exists between subsequent sets of parameters. The optimal Jastrow function u is significantly different from both the RPA function u^1 and the homogeneous form u_h of Eq. (34). The proof that this optimization succeeds in minimizing the energy expectation value may be seen in Fig. 12 (which comes from the same graphite calculation as Fig. 11). Here we plot the iterative decay of the two-body function $V^n(\mathbf{r},\mathbf{r}')$ reconstructed from the total energy coefficients determined at each iteration, according to Eq. (33). This clearly indicates the reduction to zero (within statistical noise) of the derivatives in the Euler-Lagrange equations, thus solving the energy minimization problem.

For the largest simulation cells studied ($3\times3\times3$ unit cell arrangement containing 216 electrons), we compare the optimal Jastrow correlation functions u of diamond and rhombohedral graphite. Figures 13 and 14 show the function u plotted with respect to electron separation on various line segments in the corresponding crystal structures, as already explained.

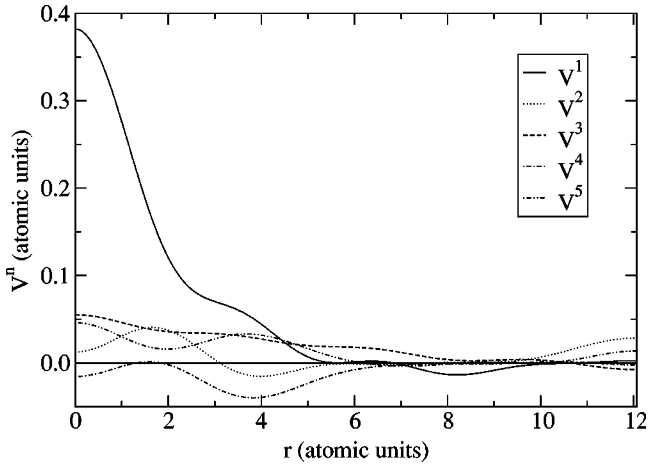

FIG. 12. The reconstructed two-body function V^n, as defined in Eq. (33), vs electron separation r on the line segment AA', for each iteration n during the optimization outlined in Figs. 7(d) and 11.

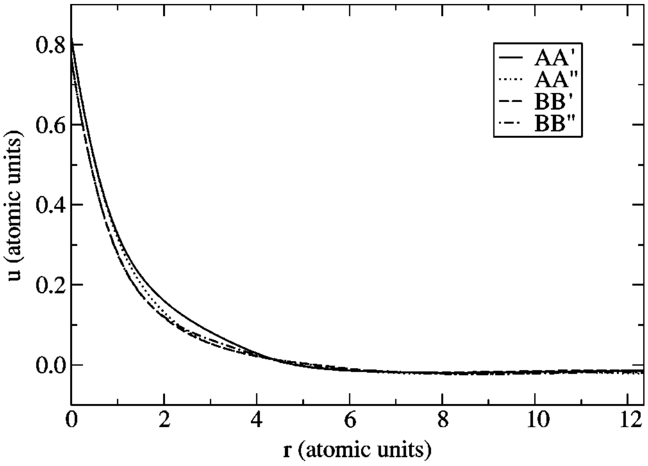

FIG. 13. The optimized diamond Jastrow correlation factor u as a function of electron separation r along the line segments indicated, for a $3\times3\times3$ simulation region.

V. DISCUSSION

A. Energies

For a direct comparison of the calculated energies of diamond and rhombohedral graphite, we should (a) make some corrections based on the trial wave functions used and (b) include finite size and zero-point phonon energy corrections for the expected energy of the real infinite solid. The corrected energies for diamond and graphite are listed in Table III.

The Jastrow factors used in both solids are comparable in their variational freedom, the only significant difference being the size of the cutoff used for χ_{1b}. However, a VMC calculation for diamond, using the same cutoff as in graphite ($G_c=3.1$ (a.u.)$^{-1}$, corresponding to 25 variational parameters), resulted in an increase in energy of only 1.0 ± 0.3 mhartree/atom for diamond. Therefore, the high Fourier coefficients of χ_{1b} contribute little to the total energy of the system. The diamond VMC energy for the $3\times3\times3$ simulation quoted in Table III was determined using $G_c=3.1$ (a.u.)$^{-1}$ as the cutoff for one-body terms.

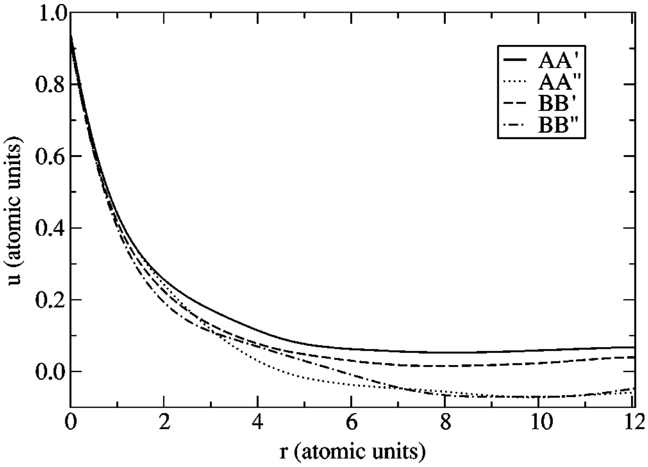

FIG. 14. The optimized rhombohedral graphite Jastrow correlation factor u as a function of electron separation r along the line segments indicated, for a $3\times3\times3$ simulation region.

TABLE III. Energies and energy corrections of diamond and rhombohedral graphite (in hartree/atom). (a) Total energy, determined by our VMC optimization method, for a $3\times3\times3$ unit cell simulation region, using wave functions of similar variational freedom (see text). (b) Finite size correction equal to the energy difference between an LDA calculation for the $3\times3\times3$ simulation region and an LDA calculation using a fully-converged k-point set (see the text). (c) Correction for the zero-point phonon energy. (Ref. 34) (d) Total energy including the corrections. The numbers in parentheses indicate the statistical error in the last digits of the corresponding energy.

	Diamond	Graphite
3×3×3	−5.712 95(14)	−5.712 91(14)
finite size correction	−0.008 99	−0.006 56
zero-point energy	0.006 65	0.006 10
total	−5.715 29(14)	−5.713 41(14)

In graphite, the exclusion of d symmetry from the basis set used to construct the LDA orbitals in the Slater determinant is energetically more important. In addition to the calculations described in Secs. IV B and IV E, we also performed calculations for graphite using a Gaussian basis set with s, p and d symmetry, and three Gaussian decays 0.22, 0.766, and 2.67. For the $3\times3\times3$ simulation, including d symmetry reduces the total VMC energy by 7.2 ± 0.3 mHartree/atom and also reduces the variance in the total energy by 16%. Use of the Ceperley-Alder exchange correlation functional to generate the single-particle orbitals in graphite, rather than the Hedin-Lundqvist form, made no difference (within statistical error) to the VMC energies, and neither did an increase in the cutoff for the Fourier space expansion of the LDA charge density from 36 to 64 Rydberg. The graphite VMC energy for the $3\times3\times3$ simulation listed in Table III was calculated using a trial wave function very similar to that used for the calculation of the corresponding diamond VMC energy. The cutoff for the one-body Jastrow factor was $G_c=3.1$ (a.u.)$^{-1}$ and the Slater determinant comprised LDA orbitals obtained using. (i) a basis set with d symmetry (as outlined above), (ii) the Ceperley-Alder exchange correlation functional, and (iii) a 64-Rydberg cutoff for the LDA charge density expansion.

We generate finite size corrections for the $3\times3\times3$ unit cell simulations of diamond and graphite, by calculating the difference in energy between a LDA calculation which uses k points compatible with periodic boundary conditions of a $3\times3\times3$ simulation region and an LDA calculation using a fully converged k-point set.[4] Comparing the change in energy between a $2\times2\times2$ calculation and a $3\times3\times3$ calculation in both diamond and graphite, using LDA and VMC methods, we see that the change in VMC energy is about 80% of the LDA energy change in diamond, and 70% in rhombohedral graphite. Perhaps more accurate estimates of the energy of the infinite solid may be obtained by an implementation of a model periodic Coulomb interaction developed recently. Tests of this approach have dramatically reduced finite size effects in the interaction energy.[35–37]

For diamond, we estimated the finite size correction to be -8.99 mhartree/atom, using a converged LDA calculation with 220 k points in the irreducible Brillouin zone. For rhombohedral graphite, incorporating d symmetry in the basis set (as described above), and using 189 k points in the LDA calculation, we found the finite size correction to be -6.56 mhartree/atom. We also include the calculated zero-point phonon energies of diamond and graphite, which are 6.65 and 6.10 mhartree/atom, respectively.[34]

Adding all these corrections to the calculated VMC energies (Table III), we estimate the energies of the infinite solids to be -5.71529 ± 0.00014 hartree/atom for diamond and -5.71341 ± 0.00014 hartree/atom for rhombohedral graphite. This appears to indicate that rhombohedral graphite is less stable than diamond. However, given the approximation of using LDA finite size corrections, we might expect a systematic error of the order of 2 mhartree/atom in each of these results. This indicates that, at the VMC level, the solids diamond and rhombohedral graphite have very similar total energies. We note that in the atomic pseudopotential used in the calculations presented here p and higher angular momentum scattering are all included in the local potential. It is possible that the use of a separate d pseudopotential might slightly affect the relative energies in both systems.

In order to determine the cohesive energy of a solid, we should subtract the energy per atom of the solid from the energy of the isolated atom, $E_c = E_a - E_s$. However, when using approximate eigenfunctions, a reasonable estimate of the cohesive energy is obtainable only by subtracting the energies estimated using similar trial wave functions. VMC energies are available for the carbon atom where the trial wave function is of the Jastrow-Slater form.[10] The orbitals of the Slater determinant are optimized using energy minimization and a sophisticated Jastrow factor is optimized using variance minimization, yielding a VMC energy for the atom of -5.4372 ± 0.0001 mhartree. Using this energy, we find the cohesive energies to be 0.2781 ± 0.0002 hartree/atom for diamond and 0.2762 ± 0.0002 hartree/atom for graphite. We regard this atomic trial wave function to be close in form and accuracy to our solid trial wave function. However, to remain consistent with the inclusion of d symmetry in the basis set of our LDA calculations, we could also refer to a multiconfiguration trial wave function for the carbon atom which includes d excitations. The VMC energy of the atom, using this wave function, is -5.45061 ± 0.00002 hartree,[10] leading to estimates of the cohesive energies of diamond and rhombohedral graphite of 0.2647 ± 0.0002 hartree/atom and 0.2628 ± 0.0002 hartree/atom, respectively. The experimental values are 0.271 hartree/atom for diamond, and 0.272 hartree/atom for graphite.[38]

Having considered the optimal homogeneous form of correlation factor for graphite, it is clear from Fig. 7 that there are only small energy gains to be made by including inhomogeneity in the correlation factor. [According to Table II this energy gain is (2.3 ± 0.2) mhartree/atom.] However, given that the optimal correlation factor for graphite is not strongly inhomogeneous (Fig. 14 and Sec. V B), it is evident that the inclusion of anisotropy in the correlation factor is energetically more important. Both the inhomogeneous and optimal homogeneous forms produce significantly lower energies than the isotropic Yukawa-style correlation factor for the larger systems.

The extra correlation energy gained using the inhomogeneous J_{ih} is not large enough to render the VMC method competitive with the DMC method for the calculation of accurate ground state energies. The decrease in the variance is $\sim5\%$, which would introduce small time savings for convergence of DMC calculations, but this saving is likely to be canceled out by the extra computational cost of evaluating J_{ih} (Sec. V C). However, to determine quantities other than total energies (e.g., the pair correlation function), may require more accurate trial wave functions. Therefore, it is important to emphasize that, whatever conclusions we would like to draw from these calculations, our principal aim has been to optimize the trial wave function for a given Hamiltonian, such that the expectation value of the total energy is minimized. This aim has been achieved for all the systems studied.

B. Correlation factors

Diamond, with a relatively homogeneous and isotropic electron charge density, exhibits an approximately homogeneous, isotropic Jastrow correlation function u (Fig. 13). At large electron separations (beyond 6 a.u.) we see that the electron correlation factor u in diamond is well approximated by homogeneous and isotropic functions, as all the curves plotted are quite similar. Only slight deviations from homogeneity exist at short and intermediate electron separations. This inhomogeneity may be seen by comparing u plotted with its fixed coordinate at different points: AA' and AA'' are quite similar at short range, but clearly distinct from BB' and BB'' in the same region. This inhomogeneity may have significant effects on the short-range pair-correlation functions calculated for diamondlike systems using VMC methods.[39–41]

Graphite, with clear regions of high electron charge density and well-defined regions of very low electron charge density between its layers, is a highly inhomogeneous and anisotropic structure. This is borne out in Fig. 14, where the function u differs considerably in various regions and in various directions. At short range, u is surprisingly homogeneous in comparison with diamond. At intermediate separations the function displays both inhomogeneous and anisotropic behaviors. Given that the layer separation in these simulations is 6.33 a.u., this indicates that inhomogeneous correlation between adjacent layers in the system is not insignificant. This may prove important for van der Waals interactions between the layers in graphite.[2]

At long range, the anisotropy of the graphite correlation factor is clearly shown in Fig. 13, where the correlation factors for electron separation vectors $\mathbf{r}'-\mathbf{r}$ lying parallel to the graphite planes (AA' and BB') are distinctly different from those with the separation vector perpendicular to the planes (AA'' and BB''). Inhomogeneity (i.e. an explicit dependence on the position $\mathbf{r}$ of the first electron) is displayed at long range in the differences between the function along AA' and BB'. We might expect this, given that the line AA' lies within the graphite planes, where the charge density is concentrated, whereas BB' lies in the very low charge density region between the planes [see Fig.2(b)]. On the other hand, when the separation vector is oriented perpendicular to the graphite planes (AA'' and BB''), inhomogeneous effects are substantially smaller at long range.

C. Computational details

In order to reduce the complexity of the physical results in Sec. IV, some computational details of the method were not discussed. We present and discuss some of these details in this section.

(1) For a system of N electrons, the computational cost of the implementation of this method of energy minimization must be evaluated. The usual computational costs of QMC calculations are not considered here, only the additional cost of the method. The evaluation of the Slater determinant still remains as a predominant factor in assessing the total cost of these calculations and improvements in that regard have recently been developed.[42]

For the optimization of a homogeneous correlation factor, such as J_{h}^*, the number of variational parameters scales as $\mathcal{O}(N)$, as does the computational cost of evaluating J_{h}^*. [We ignore the computational cost of determining the relative positions of all the electrons, which scales as $\mathcal{O}(N^2)$, since this cost is common to all types of correlation factors.] The cost, for J_{h}^*, of sampling the required expectation values to implement our energy minimization method scales as $\mathcal{O}(N^2)$, but with a very small prefactor, since for the largest, graphite $3\times3\times3$ simulation, this amounts to only 1% of the total computational cost of the calculation.

For the inhomogeneous correlation factor J_{ih}, the number of variational parameters scales as $\mathcal{O}(N^2)$. However, the prefactor on this scaling is reduced somewhat by translational symmetry. So much so, that the energy minimization method still scales quadratically with N and comprises $\sim$20% of the total calculation for the $3\times3\times3$ graphite simulation. However, the cost of the evaluation of the inhomogeneous Jastrow factor scales also as $\mathcal{O}(N^2)$ and comprises $\sim$25% of the largest calculation. For this large system of 216 electrons, the optimization of the inhomogeneous Jastrow factor takes 50% longer than the optimization of the homogeneous Jastrow factor, even though both require the same amount of VMC sampling and the same number of iterative steps for convergence.

This iterative method is trivially parallelizable. To determine the expectation values required to construct the predictor, Monte Carlo sampling may be performed independently on many workstations, and the results combined. To obtain total energies of the accuracy presented in this paper requires $\sim 10^5$ Monte Carlo samples. However, for optimization of the wave function using our method, this amount of sampling is also sufficient for accurate estimations of the expectation values required to construct the predictor (Sec. III F).

The extra time required to accumulate the various contributions to the predictor is significantly smaller than first expected, given that we must evaluate many first derivatives of some energy contributions [Eq. (28)]. There are two reasons for this: First, because we have expressed the variational components of the Jastrow factor as linear combinations of the operators $\mathcal{O}_m$ (Sec. III). These $\mathcal{O}_m$ need only be evaluated once, for each electron configuration, in order to construct both the Jastrow factor and the predictor. Second, and more importantly, since we must calculate the total energy as a sum of various contributions defined by the Hamiltonian of the system, all the energy contributions needed to construct the predictor are already available, either directly or by some simple manipulation.

The largest calculations presented in this paper were performed on a Beowulf cluster of fifteen 500 MHz dual-processor workstations. For a five iteration optimization using 10^5 samples per iteration, these calculations took about 50 h on this cluster. However, perhaps half this amount of sampling would have produced comparable results [see item (3) below]. The required memory for storing all the expectation values necessary for this calculation scales quadratically with N, but with a small prefactor. As well as all other necessary variables for the running of a VMC calculation, the storage requirements are easily within the capabilities of current workstations, at approximately 20 MB for our largest calculation.

(2) The Newton-Raphson method (see Appendix C), while quadratically convergent near a root of a multidimensional system, does possess some convergence problems far from the root. In the calculations presented in this paper, we found that below a certain minimum amount of sampling, the noise in the estimated expectation values used to construct the predictor caused divergence of the solution to Eq. (30) using the Newton-Raphson method. This problem might be solved through the use of a more robust root-finding method for the predictor function.

(3) From the analysis in Appendix G, we see that noise in the iterative method comes from the finite sampling used to estimate those expectation values (listed in Sec. III F) re-

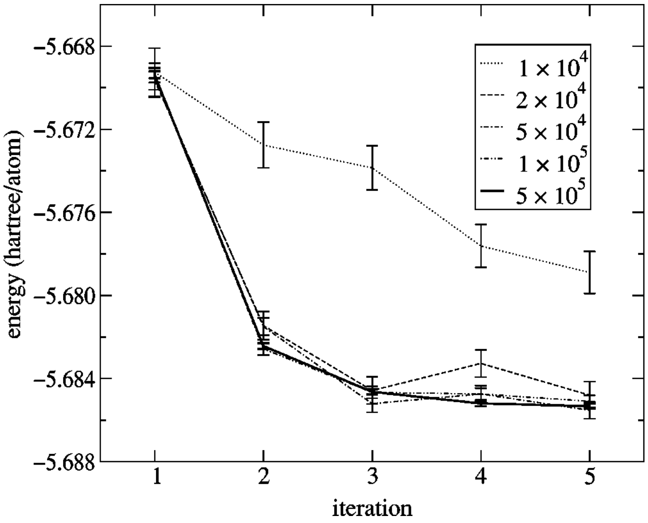

FIG. 15. Total graphite energy in hartree/atom vs iteration number for a $2\times2\times2$ simulation region. The number of Monte Carlo samples per iteration ranges from 10^4 to 5×10^5.

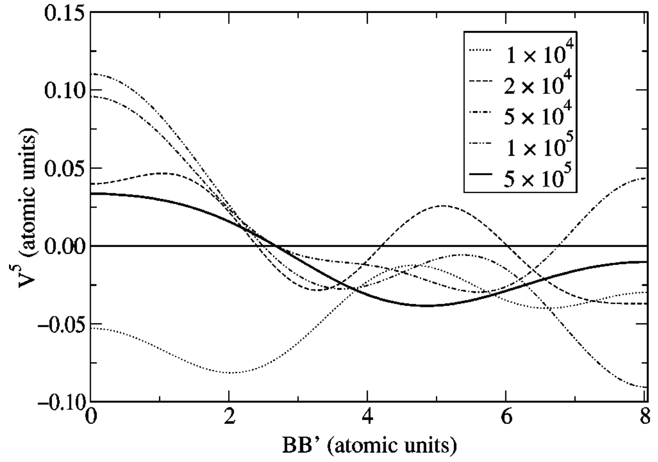

FIG. 16. The reconstructed two-body function V^5, vs electron separation r along the line segment BB' in graphite, from the fifth iteration of each of the calculations in Fig. 15.

quired to construct the predictor. The effect of this noise on the wave function accuracy is not clear. To reduce the computational cost of these calculations we would prefer to do the least amount of sampling necessary to produce the required results.

Figure 15 illustrates the effect of various amounts of sampling on the convergence of the total energy in the optimization of the $2\times2\times2$ simulation of rhombohedral graphite. At 10^4 samples per iteration, the estimation of the required expectation values, during the first iteration of our method, is too crude to produce a convergent root of the predictor using the Newton-Raphson method. This leads to a wave function, used in the second iteration, with many noisy parameters which are more difficult to optimize, as the convergence of the energy shows. However, for sampling involving 2×10^4 samples per iteration, or more, we see that the convergence of the total energy is identical (within statistical accuracy). Comparison of the optimized wave function parameters also reveals only small differences, indicating that the only means of determining the true benefits of more sampling is by examination of the fitted coefficients $V_m(\boldsymbol{\alpha}^n)$ at each iteration n.

In Fig. 16 we see, from the reconstruction $V^5(\mathbf{r},\mathbf{r}')$ of the fitted coefficients for two-body optimization at the fifth and final iteration of the method, that the most sampling (5×10^5 samples per iteration) reduces the magnitude of the Euler-Lagrange derivatives the most, indicating that these wave function parameters are the most accurate. However, for practical purposes, there is little distinction between the accuracy of the wave function once we increase the sampling beyond 2×10^4 samples per iteration. Thus the optimizations presented in Sec. IV may have used more computational time than was strictly necessary. However, more testing is required to determine the minimum amount of sampling as a function of system size.

D. Future applications

The ability to accurately represent electron correlation in systems with inhomogeneous and anisotropic electron charge densities is also pertinent for the study of finite systems such as molecules and clusters. While no work has been done to date using our method of energy minimization for inhomogeneous correlation factors in finite, polyatomic systems, the general theory of constructing the correlation factor (Sec. II) and the methodology of energy minimization (Sec. III) are still applicable here. Ultimately, what one requires is an appropriate basis set for the expansion of the correlation factor u.

For periodic solids, optimal inhomogeneous correlation factors can be used in the investigation of van der Waals interactions, and their contribution to the interlayer binding of graphite.[2] Such information may prove invaluable to the development of new exchange-correlation functionals in DFT which can accurately reproduce van der Waals energies.

VI. CONCLUSION

We have developed a generalized form of electron correlation factor for trial many-body wave functions of electrons in periodic solids. This form allows us to represent fully inhomogeneous electron correlation in real physical periodic systems. It is computationally efficient to evaluate, since the electron cusp, which we express as a homogeneous correlation factor, is separated from the fully inhomogeneous form.

We have also developed a rapidly convergent iterative method for the optimization of all variational parameters in these wave functions, minimizing the total energy of the given system. It uses the accurate techniques of quantum Monte Carlo sampling to achieve this optimization and has allowed insights into the form of many-electron correlation in systems with highly inhomogeneous charge densities.

We have demonstrated that using an optimal form of homogeneous correlation factor is appropriate for the generation of accurate energies with relatively low computational cost. In graphite, it is evident that allowing for a description of anisotropy in the correlation factor is energetically more important than representing full inhomogeneity.

Given the variational freedom to include full inhomogeneity in the correlation factor, we have found that, in diamond, the optimal correlation factor is approximately homogeneous and isotropic, with some inhomogeneity at short- and intermediate-range electron separations. This is consistent with its comparatively homogeneous and isotropic electron charge density. Graphite has an optimal correlation factor which is quite homogeneous at short-range electron separation, but is significantly inhomogeneous and anisotropic at intermediate- and long-range electron separations, as one might expect from its highly inhomogeneous and anisotropic electron charge density. Nevertheless, it is remarkable that despite the very large inhomogeneity in the electron pair-correlation function, found by previous authors,[39–41] the ideal inhomogeneous Jastrow two-body term, calculated here for diamond and graphite, displays a relatively small inhomogeneity. Whether this conclusion can be extended to other systems (e.g. involving strongly correlated d electrons) remains an open question.

ACKNOWLEDGMENTS

This work was supported by Enterprise Ireland Grant No. SC/98/748 and by the Irish Higher Education Authority. The work of S.F. at the National Microelectronics Research Centre is supported by Science Foundation Ireland.

APPENDIX A: DERIVATIVE OF OBSERVABLES WITH RESPECT TO WAVE FUNCTION PARAMETERS

The derivative of the expectation value of an observable A, with respect to a parameter $\alpha_m \in \boldsymbol{\alpha}$ of the parametrized wave function $\Psi(\boldsymbol{\alpha})$, may be written as

$$\frac{\partial\langle A\rangle}{\partial\alpha_m}=\frac{\partial}{\partial\alpha_m}\frac{\langle\Psi|A|\Psi\rangle}{\langle\Psi|\Psi\rangle}.$$

We assume that A is independent of the parameters $\boldsymbol{\alpha}$. Also, for a time independent system, we may in general express Ψ as a real function. Differentiating, we find that

$$\frac{\partial}{\partial\alpha_m}\langle\Psi|\Psi\rangle=2\langle\Psi|\mathcal{O}_m|\Psi\rangle,$$

$$\frac{\partial}{\partial\alpha_m}\langle\Psi|A|\Psi\rangle=2\langle\Psi|A\mathcal{O}_m|\Psi\rangle,$$

where we define the operator associated with variations of α_m as $\mathcal{O}_m\equiv\partial/\partial\alpha_m$, with a local value

$$\mathcal{O}_m(\mathbf{R})\equiv\frac{1}{\Psi(\mathbf{R})}\frac{\partial\Psi(\mathbf{R})}{\partial\alpha_m}=\frac{\partial}{\partial\alpha_m}\ln\Psi(\mathbf{R}),$$

for the many-body configuration $\mathbf{R}$.

Therefore, we may express the derivative of $\langle A\rangle$ as

$$\frac{1}{2}\frac{\partial\langle A\rangle}{\partial\alpha_m}=\frac{\langle\Psi|A\mathcal{O}_m|\Psi\rangle}{\langle\Psi|\Psi\rangle}-\frac{\langle\Psi|A|\Psi\rangle}{\langle\Psi|\Psi\rangle}\frac{\langle\Psi|\mathcal{O}_m|\Psi\rangle}{\langle\Psi|\Psi\rangle}$$
$$=\langle A\mathcal{O}_m\rangle-\langle A\rangle\langle\mathcal{O}_m\rangle=\langle\Delta A\Delta\mathcal{O}_m\rangle,$$

where $\Delta A(\mathbf{R})=A(\mathbf{R})-\langle A\rangle$ and $\Delta\mathcal{O}_m(\mathbf{R})=\mathcal{O}_m(\mathbf{R})-\langle\mathcal{O}_m\rangle$.

APPENDIX B: LEAST-SQUARES FITTING

Our task is to minimize the integral

$$\chi^2=\langle\Psi|\left\{\mathcal{H}-E_0-\sum_k V_k\mathcal{O}_k\right\}^2|\Psi\rangle$$

by choosing the appropriate parameters E_0 and $\{V_k\}$. We note first that, at the minimum,

$$E_0=\langle E\rangle-\sum_k V_k\langle\mathcal{O}_k\rangle,$$

where $\langle E\rangle$ is the expectation value of the total energy and $\langle\mathcal{O}_k\rangle$ is the expectation value of the operator $\mathcal{O}_k$. To fulfill the minimization, we must set all the remaining first derivatives of χ^2 to zero, i.e.,

$$\frac{\partial\chi^2}{\partial V_l}=-2\langle\Psi|\left\{\mathcal{H}-E_0-\sum_k V_k\mathcal{O}_k\right\}\mathcal{O}_l|\Psi\rangle=0$$

for each l. Upon substitution of the minimum value of E_0, this leads to

$$\langle E\mathcal{O}_l\rangle-\langle E\rangle\langle\mathcal{O}_l\rangle=\sum_k V_k[\langle\mathcal{O}_k\mathcal{O}_l\rangle-\langle\mathcal{O}_k\rangle\langle\mathcal{O}_l\rangle].$$

Since, for any operators, a and b, we may say that

$$\langle ab\rangle-\langle a\rangle\langle b\rangle=\langle\Delta a\Delta b\rangle$$

where $\Delta a=a-\langle a\rangle$, then we are lead to the conclusion that the least squares fitting is equivalent to solving the linear system

$$\sum_k V_k\langle\Delta\mathcal{O}_k\Delta\mathcal{O}_l\rangle=\langle\Delta E\Delta\mathcal{O}_l\rangle$$

for each l.

APPENDIX C: NEWTON-RAPHSON METHOD

An integral part of the iterative procedure outlined in Sec. III C is the determination of the parameters $\boldsymbol{\alpha}$ that solve the system in Eq. (30). The determination of the roots of any multidimensional function can be troublesome. In all the optimizations presented in this paper, the predictor function $V'_m(\boldsymbol{\alpha};\boldsymbol{\alpha}^0)$ is a quadratic function of $\boldsymbol{\alpha}$, whose coefficients are determined analytically, or by numerical fitting at the point $\boldsymbol{\alpha}^0$. We ignore the implicit dependence of V'_m on $\boldsymbol{\alpha}^0$ for the purpose of finding a root, and solve the system $V'_m(\boldsymbol{\alpha})=0$ for all m.

We use the Newton-Raphson method[27] to determine the roots. This is an iterative method of improving successive guesses for a root of a function. It involves the computation of the function V'_m and its Jacobian matrix of derivatives with respect to $\boldsymbol{\alpha}$, at each guess. The iterations continue until convergence of the solution is achieved within a predefined tolerance.

The success of the Newton-Raphson method for multidimensional systems relies heavily on the proximity of the initial guess to the root we seek. For this reason, the initial guess chosen is normally the parameter set of the previous iteration of the procedure outlined in Sec. III C. To find the first set of parameters, by finding the roots of the analytic predictor S_m, we require a good initial guess of the root. Since the one-body parameters $\chi_{\mathbf{G}}$ are expected to be small by construction, an initial guess of zero for all $\mathbf{G}$ was found to be sufficient to produce a convergent solution to the first application of the Newton-Raphson method.

For the two-body problem, we rescale the variables $u_{\mathbf{qq}'}$ to improve the convergence of the root-finding method. According to the RPA,[18] the long-range behavior of the u function should take the form $u(r)=1/\omega_p r$, where the plasma frequency for a homogeneous system with electron charge density n is $\omega_p=\sqrt{4\pi n}$. The charge density n is determined in the simulation region to be N/Ω, for N the number of electrons in the simulation volume Ω. Therefore, for small $\mathbf{q}$ in Fourier space, u behaves like

$$u(\mathbf{q}) = \frac{4\pi}{\Omega\omega_p q^2} = \omega_p \frac{1}{Nq^2}.$$

This indicates large relative differences between values of $u(\mathbf{q})$ for small $\mathbf{q}$. Therefore, in inhomogeneous systems, it would be appropriate to rescale the parameters $u_{\mathbf{qq}'}$ by multiplying by $N|\mathbf{q}||\mathbf{q}'|$, thus rendering all the variables of the same order of magnitude as the plasma frequency. This leads to a less pathological numerical problem for the Newton-Raphson method. Appropriate scaling must also be applied to the predictor function $V'_{\mathbf{qq}'}$ in Sec. III. For the initial analytic guess of the roots of $S_{\mathbf{qq}'}$ we use the homogeneous solution for $u_{\mathbf{qq}'}$ outlined in Eq. (31).

APPENDIX D: ONE-BODY ENERGY CONTRIBUTIONS

The variational part of Ψ associated with one-body terms is the Jastrow factor $J_{1\mathrm{b}}$ [Eq. (15)] with parameters $\boldsymbol{\chi} = \{\chi_{\mathbf{G}}\}$. We expand the energy contributions $\epsilon^{(1)}$ and $\epsilon^{(2)}$ defined in Sec. III D. For $J_{1\mathrm{b}}$, we have that

$$\begin{aligned}\epsilon^{(1)} &\equiv -\frac{1}{2}\sum_i \frac{\nabla_i^2 J_{1\mathrm{b}}}{J_{1\mathrm{b}}} \\ &= -\frac{1}{2}\sum_i (\nabla_i^2 \ln J_{1\mathrm{b}} + |\boldsymbol{\nabla}_i \ln J_{1\mathrm{b}}|^2) \\ &= \frac{1}{2}\sum_{\mathbf{G}} \chi_{\mathbf{G}} G^2 \Delta\rho_{\mathbf{G}}^* \\ &\quad + \frac{1}{2}\sum_{\mathbf{GG}'} \chi_{\mathbf{G}-\mathbf{G}'}(\mathbf{G}-\mathbf{G}')\cdot\mathbf{G}'\chi_{\mathbf{G}'}\Delta\rho_{\mathbf{G}}^* + \text{const.}\end{aligned}$$

The constant terms are not required, so we ignore them.

The energy contribution $\epsilon^{(2)}$ cannot be expanded analytically as a linear combination of the functions $\Delta\rho_{\mathbf{G}}^*$ since $\Psi' \equiv \Psi/J_{1\mathrm{b}}$ is not explicitly a function of these coordinates.

$$\epsilon^{(2)} \equiv -\sum_i \frac{\boldsymbol{\nabla}_i J_{1\mathrm{b}}}{J_{1\mathrm{b}}}\cdot\frac{\boldsymbol{\nabla}_i\Psi'}{\Psi'} = -\sum_{\mathbf{G}} \chi_{\mathbf{G}} \sum_i \boldsymbol{\nabla}_i \Delta\rho_{\mathbf{G}}^* \cdot \frac{\boldsymbol{\nabla}_i\Psi'}{\Psi'}.$$

However, $\epsilon^{(2)}$ is explicitly linear in the parameters $\boldsymbol{\chi}$, with derivatives

$$\frac{\partial\epsilon^{(2)}}{\partial\chi_{\mathbf{G}}} = -\sum_i \boldsymbol{\nabla}_i \Delta\rho_{\mathbf{G}}^* \cdot \frac{\boldsymbol{\nabla}_i\Psi'}{\Psi'},$$

which are independent of $\boldsymbol{\chi}$.

APPENDIX E: TWO-BODY ENERGY CONTRIBUTIONS

The variational part of Ψ associated with two-body energy contributions is the inhomogeneous Jastrow factor J_{ih} of Eq. (14), with parameters $\mathbf{u} = \{u_{\mathbf{qq}'}\}$. (For periodic systems we use only $u_{\mathbf{q}+\mathbf{G},\mathbf{q}+\mathbf{G}'}$.) The energy contributions $\epsilon^{(i)}$ of Sec. III D are expanded here. The contribution $\epsilon^{(1)}$ is dependent only on the form of J_{ih} and is expanded as

$$\epsilon^{(1)} \equiv -\frac{1}{2}\sum_i \frac{\nabla_i^2 J_{\mathrm{ih}}}{J_{\mathrm{ih}}} = -\frac{1}{2}\sum_i (\nabla_i^2 \ln J_{\mathrm{ih}} + |\boldsymbol{\nabla}_i \ln J_{\mathrm{ih}}|^2).$$

We find that

$$\begin{aligned}-\frac{1}{2}\sum_i \nabla_i^2 \ln J_{\mathrm{ih}} &= -\frac{1}{2}\sum_{\mathbf{qq}'} u_{\mathbf{qq}'}(q^2+q'^2)P_{\mathbf{qq}'} \\ &\quad -\sum_{\mathbf{q}}\left[\frac{N-1}{N}\sum_{\mathbf{q}'} u_{\mathbf{qq}'}q^2\langle\rho_{\mathbf{q}'}\rangle\right]\Delta\rho_{\mathbf{q}}^* \\ &\quad + \text{const.}\end{aligned}$$

We retain only the linear combination of the functions $P_{\mathbf{qq}'}$. The constant terms we may ignore, and the one-body terms (linear combinations of $\Delta\rho_{\mathbf{q}}^*$) we assume are compensated for by terms in the one-body Jastrow $J_{1\mathrm{b}}$.

If we ignore the removal of one-body terms from the function $P_{\mathbf{qq}'} = (\Delta\rho_{\mathbf{q}}\Delta\rho_{\mathbf{q}'}^*)_{[i\neq j]}$ and consider using the function $\Delta\rho_{\mathbf{q}}\Delta\rho_{\mathbf{q}'}^*$ instead, then we find that

$$\begin{aligned}-\frac{1}{2}\sum_i |\boldsymbol{\nabla}_i \ln J_{\mathrm{ih}}|^2 \\ = -2\sum_{\mathbf{qq}'}\sum_{\mathbf{kk}'} u_{\mathbf{qk}}(\mathbf{k}\cdot\mathbf{k}')u_{\mathbf{k}'\mathbf{q}'}\rho_{\mathbf{k}'-\mathbf{k}}\Delta\rho_{\mathbf{q}}\Delta\rho_{\mathbf{q}'}^* + \cdots,\end{aligned} \tag{E1}$$

where we have ignored constant and one-body terms in the expansion. The product $\rho_{\mathbf{k}'-\mathbf{k}}\Delta\rho_{\mathbf{q}}\Delta\rho_{\mathbf{q}'}^*$ contains both two- and three-body terms, since we may rewrite $\rho_{\mathbf{k}'-\mathbf{k}}$ as $\Delta\rho_{\mathbf{k}'-\mathbf{k}} + \langle\rho_{\mathbf{k}'-\mathbf{k}}\rangle$. We intend here to remove two-body fluctuations and regard three-body fluctuations as much less significant, so we retain only the two-body term[47] $\langle\rho_{\mathbf{k}'-\mathbf{k}}\rangle\Delta\rho_{\mathbf{q}}\Delta\rho_{\mathbf{q}'}^*$ from the charge fluctuation products in Eq. (E1), i.e.,

$$-2\sum_{\mathbf{qq}'}\sum_{\mathbf{kk}'} u_{\mathbf{qk}}(\mathbf{k}\cdot\mathbf{k}')u_{\mathbf{k}'\mathbf{q}'}\langle\rho_{\mathbf{k}'-\mathbf{k}}\rangle\Delta\rho_{\mathbf{q}}\Delta\rho_{\mathbf{q}'}^* + \cdots.$$

Now we make the assumption that removing one-body terms from this expression , i.e., replacing $\Delta\rho_{\mathbf{q}}\Delta\rho_{\mathbf{q}'}^*$ with $P_{\mathbf{qq}'}$ is a good approximation, and obtain the expression for $v_{\mathbf{qq}'}^{(1)}(\mathbf{u})$ given in Sec. III E 2.

The contribution $\epsilon^{(2)}$ cannot be expanded analytically in the basis of fluctuation functions $P_{\mathbf{qq}'}$, However, it is clear that $\epsilon^{(2)}$ is linear in $\mathbf{u}$, since, for $\Psi' \equiv \Psi/J_{\mathrm{ih}}$,

$$\epsilon^{(2)} = \sum_{\mathbf{qq}'} u_{\mathbf{qq}'}\sum_i \boldsymbol{\nabla}_i P_{\mathbf{qq}'}\cdot\frac{\boldsymbol{\nabla}_i\Psi'}{\Psi'},$$

and has first derivatives

$$\frac{\partial\epsilon^{(2)}}{\partial u_{\mathbf{qq}'}} = \sum_i \boldsymbol{\nabla}_i P_{\mathbf{qq}'}\cdot\frac{\boldsymbol{\nabla}_i\Psi'}{\Psi'},$$

which are independent of the parameters $\mathbf{u}$.

The final energy contribution $\epsilon^{(3)}$ we attempt to express analytically in terms of the two-body fluctuation functions $P_{\mathbf{q}\mathbf{q}'}$. If $\Psi' \approx D$, the LDA Slater determinant, then the sum of contributions from the external potential and the kinetic energy term $-(1/2)\Sigma_i \nabla_i^2 \Psi'/\Psi'$ is a one-body contribution defined by the Kohn-Sham Hamiltonian,[25] since

$$\sum_i \left[-\frac{1}{2}\frac{\nabla_i^2 D}{D} + V_{\mathrm{ext}}(\mathbf{r}_i) \right] = \sum_i \left[\epsilon_i^{\mathrm{KS}} - V_{\mathrm{H}}(\mathbf{r}_i) - V_{\mathrm{xc}}(\mathbf{r}_i) \right],$$

where ϵ_i^{KS} are the Kohn-Sham eigenvalues, V_{H} is the Hartree potential and V_{xc} is the exchange and correlation potential. Therefore, the two-body contribution of these terms is approximately zero. We are left with the contribution of the electron-electron interaction V.

For a two-body potential $V(\mathbf{r},\mathbf{r}')$, we may expand the sum over electron pairs as

$$\sum_{i<j} V(\mathbf{r}_i,\mathbf{r}_j) = \frac{1}{2}\sum_{\mathbf{q}\mathbf{q}'} V_{\mathbf{q}\mathbf{q}'} \sum_{i \neq j} e^{-i\mathbf{q}\cdot\mathbf{r}_i} e^{i\mathbf{q}'\cdot\mathbf{r}_j}$$
$$= \frac{1}{2}\sum_{\mathbf{q}} V_{\mathbf{q}\mathbf{q}'} (\rho_{\mathbf{q}}\rho_{\mathbf{q}'}^{*})_{[i \neq j]},$$

for Fourier coefficients $V_{\mathbf{q}\mathbf{q}'}$. In terms of the fluctuation functions $P_{\mathbf{q}\mathbf{q}'}$, this may be rewritten as

$$\sum_{i<j} V(\mathbf{r}_i,\mathbf{r}_j) = \frac{1}{2}\sum_{\mathbf{q}\mathbf{q}'} V_{\mathbf{q}\mathbf{q}'} P_{\mathbf{q}\mathbf{q}'}$$
$$-\sum_{\mathbf{q}} \left[\frac{N-1}{N} \sum_{\mathbf{q}'} V_{\mathbf{q}\mathbf{q}'} \langle \rho_{\mathbf{q}'} \rangle \right] \Delta\rho_{\mathbf{q}}^{*} + \text{ const.}$$

[We assume that $V(\mathbf{r},\mathbf{r}')$ possesses exchange symmetry, so that $V_{\mathbf{q}\mathbf{q}'} = V_{-\mathbf{q}'-\mathbf{q}}$.] Therefore, both one- and two-body fluctuations arise from a two-body potential in the "charge fluctuation" coordinate system. Note that the one-body fluctuations are expressible in terms of the Hartree potential, since

$$V_{\mathrm{H}}(\mathbf{q}) = \sum_{\mathbf{q}'} V_{\mathbf{q}\mathbf{q}'} \langle \rho_{\mathbf{q}'} \rangle.$$

If the two-body potential is homogeneous, i.e., $V(\mathbf{r},\mathbf{r}') = V(\mathbf{r}-\mathbf{r}')$, then we may simplify the fluctuations since $V_{\mathbf{q}\mathbf{q}'} = V_{\mathbf{q}}^{*}\delta(\mathbf{q}-\mathbf{q}')$, where the Fourier transform of the homogeneous function $V(\mathbf{r})$ is

$$V_{\mathbf{q}} \equiv \frac{1}{\Omega}\int_{\Omega} V(\mathbf{r}) e^{-i\mathbf{q}\cdot\mathbf{r}} d\mathbf{r},$$

for a system volume Ω. If $V(\mathbf{r}) = V(-\mathbf{r})$ then $V_{\mathbf{q}}^{*} = V_{\mathbf{q}}$. Therefore,

$$\sum_{i<j} V(|\mathbf{r}_i - \mathbf{r}_j|) = \frac{1}{2}\sum_{\mathbf{q}} V_{\mathbf{q}} P_{\mathbf{q}\mathbf{q}}$$
$$-\frac{N-1}{N}\sum_{\mathbf{q}} V_{\mathbf{q}} \langle \rho_{\mathbf{q}} \rangle \Delta\rho_{\mathbf{q}}^{*} + \text{ const.} \tag{E2}$$

Again we ignore the one-body and constant terms in this context.

APPENDIX F: CONSEQUENCES OF USING THE SHORT-RANGE JASTROW FACTOR

The short range Jastrow factor J_{sr} is constructed from the pseudointeraction V_{ps} in the following way. For the isolated two-electron scattering problem, we may find an eigenstate ψ_0 of the true two-electron Hamiltonian h_0 for a given energy eigenvalue ϵ. Upon replacing the true Coulomb interaction V with a generated pseudointeraction V_{ps}, we construct the modified Hamiltonian h_{ps} with an eigenstate ψ_{ps} corresponding to ϵ. We construct J_{sr} such that

$$\psi_0 = J_{\mathrm{sr}} \psi_{\mathrm{ps}}.$$

Now, in a many-electron environment, we know that using J_{sr} allows for a good approximation of short-range correlations.[24] We might imagine that for a many-electron system with Hamiltonian $\mathcal{H}$ and many-electron trial wave function $J_{\mathrm{sr}}\Psi_{\mathrm{ps}}$, the true energy eigenvalue may be well approximated by

$$E \approx \frac{\mathcal{H} J_{\mathrm{sr}} \Psi_{\mathrm{ps}}}{J_{\mathrm{sr}} \Psi_{\mathrm{ps}}} \approx \frac{\mathcal{H}_{\mathrm{ps}} \Psi_{\mathrm{ps}}}{\Psi_{\mathrm{ps}}},$$

where $\mathcal{H}_{\mathrm{ps}} = \mathcal{H} - \Sigma_{i<j}[V(r_{ij}) - V_{\mathrm{ps}}(r_{ij})]$. Now, disguising the true interaction V with $V_{\mathrm{ps}} + (V - V_{\mathrm{ps}})$, and given the transferability of V_{ps} over a wide range of energies, we see that

$$\sum_{i<j} V_{\mathrm{ps}}(r_{ij}) \approx -\frac{1}{2}\sum_i \frac{\nabla_i^2 J_{\mathrm{sr}}}{J_{\mathrm{sr}}} - \sum_i \frac{\boldsymbol{\nabla}_i J_{\mathrm{sr}}}{J_{\mathrm{sr}}} \cdot \frac{\boldsymbol{\nabla}_i \Psi_{\mathrm{ps}}}{\Psi_{\mathrm{ps}}}$$
$$+ \sum_{i<j} V(r_{ij}).$$

Dividing two-body correlation into short-range and inhomogeneous terms, we use a Jastrow factor of the form $J_{\mathrm{sr}}J_{\mathrm{ih}}$. The local energy determined using the trial wave function $\Psi = J_{\mathrm{sr}} J_{\mathrm{ih}} \Psi'$, where $\Psi' \equiv \Psi/(J_{\mathrm{sr}} J_{\mathrm{ih}})$, may be expanded as

$$E = -\frac{1}{2}\sum_i \frac{\nabla_i^2 J_{\mathrm{sr}}}{J_{\mathrm{sr}}} - \sum_i \frac{\boldsymbol{\nabla}_i J_{\mathrm{sr}}}{J_{\mathrm{sr}}} \cdot \frac{\boldsymbol{\nabla}_i \Psi/J_{\mathrm{sr}}}{\Psi/J_{\mathrm{sr}}} + \sum_{i<j} V(r_{ij})$$
$$-\frac{1}{2}\sum_i \frac{\nabla_i^2 J_{\mathrm{ih}}}{J_{\mathrm{ih}}} - \sum_i \frac{\boldsymbol{\nabla}_i J_{\mathrm{ih}}}{J_{\mathrm{ih}}} \cdot \frac{\boldsymbol{\nabla}_i \Psi'}{\Psi'} - \frac{1}{2}\sum_i \frac{\nabla_i^2 \Psi'}{\Psi'}$$
$$+ \sum_i V_{\mathrm{ext}}(\mathbf{r}_i)$$

$$\approx \sum_{i<j} V_{\mathrm{ps}}(r_{ij}) - \frac{1}{2}\sum_i \frac{\nabla_i^2 J_{\mathrm{ih}}}{J_{\mathrm{ih}}} - \sum_i \frac{\nabla_i J_{\mathrm{ih}}}{J_{\mathrm{ih}}} \cdot \frac{\nabla_i \Psi'}{\Psi'}$$

$$-\frac{1}{2}\sum_i \frac{\nabla_i^2 \Psi'}{\Psi'} + \sum_i V_{\mathrm{ext}}(\mathbf{r}_i).$$

For this reason, we use V_{ps} in the expansion of the local energy for J_{ih} to implicitly include the short-range Jastrow factor J_{sr}.

APPENDIX G: CONVERGENCE OF THE METHOD

The convergence criterion $V_m(\boldsymbol{\alpha})=0$, is numerically never exactly achieved. Given that the predictor V_m' contains some terms determined by statistical fitting, finite sampling errors exist, and this noise is passed on to the fitted parameters in V_m' from Eqs. (22) and (27).

We consider $\mathbf{V}_n=\{V_m(\boldsymbol{\alpha}^n)\}$, the fitted coefficients of the local energy at the n^{th} iteration of the method. The method may be regarded as an iterative map $\mathbf{M}$, such that the coefficients are determined via $\mathbf{V}_{n+1}=\mathbf{M}(\mathbf{V}_n)$. There are two sources of noise in $\mathbf{V}_{n+1}$: (i) noise inherited from $\mathbf{V}_n$, which produced the parameters $\boldsymbol{\alpha}^{n+1}$, which were used to construct the wave function $\Psi(\boldsymbol{\alpha}^{n+1})$, with which we evaluated the expectation values used to calculate $\mathbf{V}_{n+1}$; and (ii) finite sampling noise in the evaluation of the expectation values in Eqs. (22) and (27) via Monte Carlo sampling. Therefore, we associate a set of variances $\boldsymbol{\sigma}_n^2=\{\sigma_m^2$ for each m at step $n\}$, arising from these two sources of noise, to the coefficients $\mathbf{V}_n$. The variance due to finite sampling alone, at each step n, is $\mathbf{s}_n^2$ and we use the initial condition $\boldsymbol{\sigma}_1^2=\mathbf{s}_1^2$. This implies that the variance obeys the following iterative map:

$$\boldsymbol{\sigma}_{n+1}^2 \approx \mathbf{s}_{n+1}^2 + |\boldsymbol{\lambda}_n|^2 \boldsymbol{\sigma}_n^2,$$

where $\boldsymbol{\lambda}_n=\nabla_{\mathbf{V}}\mathbf{M}(\mathbf{V}_n)$. For a convergent map $\mathbf{M}$, we are guaranteed that $|\boldsymbol{\lambda}_n|^2<1$ at convergence. $|\boldsymbol{\lambda}_n|$ is a measure of the convergence rate of the map $\mathbf{M}$, with $|\boldsymbol{\lambda}_n|\approx 0$ implying fast convergence and $|\boldsymbol{\lambda}_n|\approx 1$ implying slow convergence. Note that if $|\boldsymbol{\lambda}_n|\geqslant 1$ the map is divergent.

Therefore, if we regard $\mathbf{s}_n\approx\mathbf{s}$ and $\boldsymbol{\lambda}_n\approx\boldsymbol{\lambda}$ for all n, for constants $\mathbf{s}$ and $\boldsymbol{\lambda}$, then the converged value of the variance in the fitted coefficients is

$$\boldsymbol{\sigma}_*^2 = \frac{1}{1-|\boldsymbol{\lambda}|^2}\mathbf{s}^2. \tag{G1}$$

We conclude that, given a convergent method $\mathbf{M}$, the presence of statistical noise does not lead to successively more ill-determined parameters $\boldsymbol{\alpha}$, since their variance is also convergent.

Note that Eq. (G1) indicates that the variance in the fitted coefficients $V_m(\boldsymbol{\alpha})$ is always greater than or equal to the variance estimated using finite sampling. However, our implementation of the method represented by the map $\mathbf{M}$ indicates that $|\boldsymbol{\lambda}|\ll 1$, since we find that the majority of the coefficients $V_m(\boldsymbol{\alpha})$ ultimately end up with magnitudes approximately equal to their finite sampling errors, signifying that statistically they are zero.

Therefore, the final conclusion to be drawn from Eq. (G1) is that the accuracy of our optimization method depends ultimately on the finite sampling error. Therefore, increasing the computational workload, by increasing the amount of sampling, will result in more accurate optimizations of the wave function. This is demonstrated by the results in Sec. V.

---

*Present address: Institute of Metallurgy and Materials Science, The University of Birmingham, Edgbaston, Birmingham, B15 2TT, United Kingdom.

[1] P. Hohenberg and W. Kohn, Phys. Rev. **136**, B864 (1964).

[2] D. Prendergast and S. Fahy (unpublished).

[3] W.M.C. Foulkes, L. Mitas, R.J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).

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

[5] E. Feenberg, *Theory of Quantum Fluids* (Academic, New York, 1969).

[6] E. Krotscheck, W. Kohn, and G.-X. Qian, Phys. Rev. B **32**, 5693 (1985).

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

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

[9] L. Mitas, E.L. Shirley, and D.M. Ceperley, J. Chem. Phys. **95**, 3467 (1991).

[10] C. Filippi and S. Fahy, J. Chem. Phys. **112**, 3523 (2000).

[11] F. Schautz and S. Fahy, J. Chem. Phys. **116**, 3533 (2002).

[12] M.P. Nightingale and V. Melik-Alaverdian, Phys. Rev. Lett. **87**, 043401 (2001).

[13] W.L. McMillan, Phys. Rev. **138**, A442 (1965).

[14] A. Harju, B. Barbiellini, S. Siljamaki, R.M. Nieminen, and G. Ortiz, Phys. Rev. Lett. **79**, 1173 (1997).

[15] X. Lin, H. Zhang, and A.M. Rappe, J. Chem. Phys. **112**, 2650 (2000).

[16] M. Snajdr, J.R. Dwyer, and S.M. Rothstein, J. Chem. Phys. **111**, 9971 (1999).

[17] M. Snajdr and S.M. Rothstein, J. Chem. Phys. **112**, 4935 (2000).

[18] D. Bohm and D. Pines, Phys. Rev. **92**, 609 (1953).

[19] D. Bevan, Master's thesis, National University of Ireland, Cork, Ireland, 1997.

[20] R. Gaudoin, M. Nekovee, W.M.C. Foulkes, R.J. Needs, and G. Rajagopal, Phys. Rev. B **63**, 115115 (2001).

[21] A. Malatesta, S. Fahy, and G.B. Bachelet, Phys. Rev. B **56**, 12 201 (1997).

[22] T. Kato, Commun. Pure Appl. Math. **10**, 151 (1957).

[23] J. D. Morgan III, in *Numerical Determination of the Electronic Structure of Atoms, Diatomic and Polyatomic Molecules*, edited by M. Defranceschi and J. Delhalle (Kluwer, Dordrecht, 1989), pp. 49–84.

[24] D. Prendergast, M. Nolan, C. Filippi, S. Fahy, and J.C. Greer, J. Chem. Phys. **115**, 1626 (2001).

[25] W. Kohn and L.J. Sham, Phys. Rev. **140**, 1133 (1965).

[26] S. Fahy, in *Quantum Monte Carlo Methods in Physics and Chemistry*, edited by M. P. Nightingale and C. J. Umrigar (Kluwer, Dordrecht, 1999), pp. 101–127.

[27] W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes (FORTRAN Version)* (Cambridge University Press, Cambridge, 1990).

[28] J. Donohue, *The Structures of Elements* (Wiley, New York, 1974).

[29] C.T. Chan, D. Vanderbilt, and S.G. Louie, Phys. Rev. B **33**, 2455 (1986).

[30] D. Ceperley, Phys. Rev. B **18**, 3126 (1978).

[31] L. Hedin and B.I. Lundqvist, J. Phys. C **4**, 2064 (1971).

[32] 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).

[33] L. Mitas, in *Computer Simulation Studies in Condensed-Matter Physics V*, edited by D. P. Landau, K. K. Mon, and H. B. Schuttler (Springer, Berlin, 1993).

[34] M.T. Yin and M.L. Cohen, Phys. Rev. B **29**, 6996 (1984); the zero-point phonon energy quoted for graphite was calculated for the Bernal graphite structure, but we assume that this is a good approximation to the zero-point energy of the rhombohedral structure studied in this paper.

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

[36] A.J. Williamson, G. Rajagopal, R.J. Needs, L.M. Fraser, W.M.C. Foulkes, Y. Wang, and M.-Y. Chou, Phys. Rev. B **55**, R4851 (1997).

[37] P.R.C. Kent, R.Q. Hood, A.J. Williamson, R.J. Needs, W.M.C. Foulkes, and G. Rajagopal, Phys. Rev. B **59**, 1917 (1999).

[38] L. Brewer, Lawrence Berkeley Laboratory Report No. LB-3720 (unpublished).

[39] S. Fahy, X.W. Wang, and S.G. Louie, Phys. Rev. Lett. **65**, 1478 (1990).

[40] R.Q. Hood, M.Y. Chou, A.J. Williamson, G. Rajagopal, R.J. Needs, and W.M.C. Foulkes, Phys. Rev. Lett. **78**, 3350 (1997).

[41] R.Q. Hood, M.Y. Chou, A.J. Williamson, G. Rajagopal, and R.J. Needs, Phys. Rev. B **57**, 8972 (1998).

[42] A.J. Williamson, R.Q. Hood, and J.C. Grossman, Phys. Rev. Lett. **87**, 246406 (2001).

[43] D. Bressanini, G. Morosi, and M. Mella, J. Chem. Phys. **116**, 5345 (2002).

[44] This may leave the variance minimization susceptible to instability due to the large cost associated with outliers (energies far from the constant value used in the fit). Recently, robust estimation methods have been discussed to improve the convergence and stability of variance minimization by fitting energy fluctuations, assuming that they are not necessarily normally distributed (Ref. 43).

[45] If the matrix $\langle \Delta\mathcal{O}_m \Delta\mathcal{O}_n \rangle$ is a singular, or at least numerically so (i. e., has a very small eigenvalue), a singular value decomposition is recommended to set the appropriate fitting coefficients V_l to zero. Such problems do not render the wave function parameters α_l ill conditioned, as the small size of the corresponding expectation values $\langle \Delta\mathcal{O}_l^2 \rangle$ implies that changes in α_l have little effect on the total energy.

[46] Recent tests of the iterative method (Ref. 2) have shown that the inclusion of covariance terms between one- and two-body operators from all k points (not just $\mathbf{k}=\mathbf{0}$) improves the convergence of the method and the stability of the optimized parameters, particularly for systems sensitive to small changes in the electron density, such as graphite.

[47] This approximation is related to the RPA of Ref. 18 and is used in the work of Refs. 19, 20, and 30.