PHYSICAL REVIEW E **86**, 036703 (2012)

Framework for constructing generic Jastrow correlation factors

P. López Ríos,[1] P. Seth,[1] N. D. Drummond,[2] and R. J. Needs[1]
[1]*Theory of Condensed Matter Group, Cavendish Laboratory, University of Cambridge, J. J. Thomson Avenue, Cambridge CB3 0HE, United Kingdom*
[2]*Department of Physics, Lancaster University, Lancaster LA1 4YB, United Kingdom*

(Received 13 July 2012; published 27 September 2012)

We have developed a flexible framework for constructing Jastrow factors which allows for the introduction of terms involving arbitrary numbers of particles. The use of various three- and four-body Jastrow terms in quantum Monte Carlo calculations is investigated, including a four-body van der Waals-like term, and anisotropic terms. We have tested these Jastrow factors on one- and two-dimensional homogeneous electron gases, the Be, B, and O atoms, and the BeH, H_2O, N_2, and H_2 molecules. Our optimized Jastrow factors retrieve more than 90% of the fixed-node diffusion Monte Carlo correlation energy in variational Monte Carlo for each system studied.

DOI: 10.1103/PhysRevE.86.036703 PACS number(s): 02.70.Ss, 31.15.V−, 71.10.−w, 71.15.−m

I. INTRODUCTION

The variational and diffusion quantum Monte Carlo methods (VMC and DMC) are zero-temperature stochastic techniques for evaluating the expectation values of time-independent operators [1–3]. These methods are particularly well suited for calculating the ground-state energies of large assemblies of interacting quantum particles. The central object is an approximate trial wave function whose accuracy controls the computed energy and the intrinsic statistical fluctuations in the calculations. It is therefore of particular importance to develop accurate trial wave functions.

Expectation values in VMC are evaluated using importance-sampled Monte Carlo integration. In DMC the ground state is projected out by evolving the Schrödinger equation in imaginary time. Such projector methods suffer from a fermion sign problem, in which errors in the propagation increase exponentially in imaginary time as the algorithm amplifies any spurious component of the lower-energy bosonic state. This problem may be evaded in DMC by employing the fixed-node approximation [4], in which the nodal surface is fixed to be that of a suitable trial wave function. The resulting DMC energy is greater than or equal to the exact energy and less than or equal to the VMC energy computed with the same trial wave function. The VMC energy depends on the entire trial wave function, but the DMC energy depends only on the nodal surface of the trial wave function.

One of the appealing features of VMC and DMC is that virtually any form of trial wave function can be used. The main criteria are that the wave function must obey the correct symmetry under particle exchange, it should be flexible enough to describe the system of interest, and it should be possible to evaluate it rapidly. The analytic properties and normalizability of the trial wave function must be such that the energy expectation value is well defined. The simplest fermionic wave function is a Slater determinant, which describes exchange but not correlation. Multideterminant wave functions, pairing wave functions such as geminals [5], and backflow transformations [6] can also be used. The most fruitful method of going beyond the Slater determinant is, however, to multiply it by a Jastrow factor [1], which leads to the Slater-Jastrow wave function. The Jastrow factor is normally chosen to depend on the interparticle separations, which introduces correlation into the wave function. The introduction of a Jastrow factor often leads to the recovery of 80% or more of the correlation energy of electronic systems [7].

The Jastrow factor is chosen to be everywhere positive and symmetric with respect to the exchange of identical particles in order to maintain the nodal surface defined by the rest of the wave function. One of the features of the Jastrow factor is that it can conveniently be used to enforce the Kato cusp conditions [8], which determine the behavior of the wave function when two charged particles approach one another. Enforcing the Kato cusp conditions does not necessarily improve the variational energy, but the reduction in the statistical fluctuations in the energy is often very important.

DMC can be viewed as VMC with a perfect Jastrow factor, but improving the Jastrow factor can improve DMC calculations in several ways. The DMC algorithm is subject to time-step errors and to (normally very small) population-control errors [9] that are reduced by improving the trial wave function. Evaluating expectation values of operators that do not commute with the Hamiltonian is not straightforward in DMC, but using highly accurate trial wave functions helps in achieving more accurate results. Similar considerations apply when using nonlocal pseudopotentials, which involves making approximations that are ameliorated by improving the trial wave function [10,11]. As the fundamental limitation on the accuracy of DMC is the quality of the nodal surface, it is desirable to use trial wave functions with optimizable nodal surfaces as afforded by, for example, multideterminant wave functions and backflow transformations. A good Jastrow factor can account for the bulk of the dynamical correlation energy, which allows the optimization of parameters that affect the nodal surface to achieve a better nodal surface.

Here we introduce a highly flexible form of Jastrow factor which allows for the introduction of a variety of terms involving arbitrary numbers of particles. Our main motivation is to be able to implement quickly different functional forms and explore the importance of different correlations in any physical system we study.

Jastrow factors correlating several electrons have been used in earlier calculations, such as those of Refs. [12–15]. We study the effects of various three-body Jastrow terms and introduce a four-body van der Waals-like term. We also construct anisotropic Jastrow factors that can capture the natural symmetries of a system. We have successfully applied these Jastrow factors to a variety of systems, and we report results for the one- and two-dimensional homogeneous electron gases, the Be, B, and O atoms, and the BeH, N_2, H_2O, and H_2 molecules. Our VMC and DMC calculations were performed using the CASINO package [16]. Hartree atomic units are used throughout ($\hbar = |e| = m_e = 4\pi\epsilon_0 = 1$). The structure of the paper is as follows. We describe the form and properties of the general Jastrow factor in Sec. II. Specific examples of the construction of Jastrow terms are given in Sec. III, and results obtained using them are presented in Sec. IV. Finally we draw our conclusions in Sec. V. Implementation details are given in Appendix A, and an example can be found in Appendix B. We report only summaries of our data in this paper [17].

II. CONSTRUCTION OF A GENERIC JASTROW FACTOR

Quantum Monte Carlo (QMC) methods can be applied to systems which can be generically described as an ensemble of N quantum particles and M sources of external potential. The most common type of QMC simulations are electronic calculations, where the quantum particles are electrons and the sources of external potential are fixed nuclei (or pseudopotentials). For simplicity, we refer to quantum particles as electrons and to external potentials as nuclei in the rest of this paper. Our Jastrow factor is applicable to other types of quantum particles and external potentials.

Any trial wave function can be written in the form

$$\Psi_{\mathrm{T}}(\mathbf{R}) = e^{J(\mathbf{R})}\Psi_{\mathrm{S}}(\mathbf{R}), \tag{1}$$

where $\Psi_{\mathrm{S}}(\mathbf{R})$ is the part of the wave function that imposes the symmetry and boundary conditions, and $e^{J(\mathbf{R})}$ is a Jastrow correlation factor which is constrained so that the symmetry and boundary properties of $\Psi_{\mathrm{S}}(\mathbf{R})$ are transferred unmodified to $\Psi_{\mathrm{T}}(\mathbf{R})$. In the single-determinant (SD) Slater-Jastrow wave function, $\Psi_{\mathrm{S}}(\mathbf{R})$ is a Slater determinant. There are various alternatives to Slater determinants, with different advantages and disadvantages depending on the system.

Typically $J(\mathbf{R})$ is constructed as a sum of terms, e.g.,

$$J(\mathbf{R}) = J_{\mathrm{e-e}}(\mathbf{R}) + J_{\mathrm{e-n}}(\mathbf{R}) + J_{\mathrm{e-e-n}}(\mathbf{R}) + \cdots , \tag{2}$$

where "e–e" stands for "electron-electron," "e–n" for "electron–nucleus," etc. Each of these terms involves different numbers of electrons n and nuclei m. We shall refer to n and m as the electronic and nuclear ranks of a term, respectively, which are constrained to satisfy $n + m \geqslant 2, n \geqslant 1$, and $m \geqslant 0$. We have designed a generic Jastrow term of selectable ranks, $J_{n,m}(\mathbf{R})$, such that the total Jastrow factor is constructed as the exponential of a sum of one or more terms of the desired ranks. In this notation, $J_{\mathrm{e-e}} \equiv J_{2,0}$, $J_{\mathrm{e-n}} \equiv J_{1,1}$, etc.

The function $J_{n,m}(\mathbf{R})$ is a sum over all sets of n electrons and m nuclei in the system of a parameterized function of the e–e and e–n relative position vectors within each such set. While alternatives exist, a natural way of parameterizing this function for arbitrary values of n and m (implying an arbitrary number of variables in the function) is to expand it in products of functions of the individual e–e and e–n vectors. Thus, we construct our Jastrow factor using pairwise objects as building blocks, and in what follows we describe these objects and derive the properties of $J_{n,m}(\mathbf{R})$ that follow from those of the pairwise objects.

We name the e–e functions used in the expansion $\Phi_\nu^P(\mathbf{r})$, where $\mathbf{r}$ is the relevant e–e relative position vector, ν is the index of the function within a chosen basis of functions, and P is the e–e dependency index, which allows the use of different optimizable parameters, if present, for parallel- and antiparallel-spin electron pairs, for example. Similarly, the e–n basis functions are $\Theta_\mu^S(\mathbf{r})$, where $\mathbf{r}$ is the relevant e–n relative position vector, μ is the index of the function within the chosen basis set, and S is the e–n dependency index of the basis set, which allows the use of different parameters for up- and down-spin electrons around a given nucleus, or for different atoms, for example. In the case of nonelectronic systems, e–e and e–n dependency indices are used to distinguish between particle types and spins.

We introduce a compact notation for defining $J_{n,m}(\mathbf{R})$. We represent the n electronic indices by the integer vector $\mathbf{i} = \{i_1,i_2,\ldots,i_n\}$, each of whose components takes a distinct value between 1 and N, and the m nuclear indices by the integer vector $\mathbf{I} = \{I_1,I_2,\ldots,I_m\}$, each of whose components takes a distinct value between 1 and M. For each term in the Jastrow factor we define the e–e and e–n dependency matrices $\underline{\mathrm{P}}$ and $\underline{\mathrm{S}}$ of respective sizes $N \times N$ and $N \times M$ containing the dependency indices P_{ij} and S_{iI} for each e–e and e–n pair. The components of $\underline{\mathrm{P}}$ and $\underline{\mathrm{S}}$ can be made equal depending on the symmetries of the system, including particle distinguishability and geometrical symmetries which make different nuclei equivalent.

Likewise, it is convenient to use matrices to represent the basis functions involved in the Jastrow factor term. For e–e basis functions, each row and column of the $n \times n$ matrix $\underline{\Phi}$ corresponds to an electron:

$$\underline{\Phi}_{\underline{\nu}}^{\underline{\mathrm{P}}}(\mathbf{i}) = \begin{pmatrix} 0 & \Phi_{\nu_{i_1 i_2}}^{P_{i_1 i_2}}(\mathbf{r}_{i_1 i_2}) & \cdots & \Phi_{\nu_{i_1 i_n}}^{P_{i_1 i_n}}(\mathbf{r}_{i_1 i_n}) \\ \Phi_{\nu_{i_1 i_2}}^{P_{i_1 i_2}}(\mathbf{r}_{i_2 i_1}) & 0 & \cdots & \Phi_{\nu_{i_2 i_n}}^{P_{i_2 i_n}}(\mathbf{r}_{i_2 i_n}) \\ \vdots & \vdots & \ddots & \vdots \\ \Phi_{\nu_{i_1 i_n}}^{P_{i_1 i_n}}(\mathbf{r}_{i_n i_1}) & \Phi_{\nu_{i_2 i_n}}^{P_{i_2 i_n}}(\mathbf{r}_{i_n i_2}) & \cdots & 0 \end{pmatrix}. \tag{3}$$

We refer to the $n \times n$ matrix formed by the e–e dependency indices $\{P_{i_\alpha i_\beta}\}_{\alpha,\beta=1,\ldots,n}$ as $\underline{\mathrm{P}}(\mathbf{i})$. Both $\underline{\mathrm{P}}(\mathbf{i})$ and the $n \times n$ matrix of e–e expansion indices $\underline{\nu}$ are defined to be symmetric, and this fact has been used in Eq. (3). Noting that $\mathbf{r}_{ji} = -\mathbf{r}_{ij}$, and restricting the e–e functions to be either symmetric or antisymmetric about the origin, one finds in Eq. (3) that matrix $\underline{\Phi}$ is symmetric, antisymmetric, or asymmetric depending on whether the functions in the basis set are all symmetric, all antisymmetric, or both types are present, respectively.

For e–n basis functions each row of the $n \times m$ matrix $\underline{\Theta}$ corresponds to an electron and each column to a nucleus:

$$\underline{\Theta}^{\underline{\mathrm{S}}}_{\underline{\mu}}(\mathbf{i},\mathbf{I}) = \begin{pmatrix} \Theta^{S_{i_1 I_1}}_{\mu_{i_1 I_1}}(\mathbf{r}_{i_1 I_1}) & \cdots & \Theta^{S_{i_1 I_m}}_{\mu_{i_1 I_m}}(\mathbf{r}_{i_1 I_m}) \\ \vdots & & \vdots \\ \Theta^{S_{i_n I_1}}_{\mu_{i_n I_1}}(\mathbf{r}_{i_n I_1}) & \cdots & \Theta^{S_{i_n I_m}}_{\mu_{i_n I_m}}(\mathbf{r}_{i_n I_m}) \end{pmatrix}. \tag{4}$$

We refer to the $n \times m$ matrix formed by the e–n dependency indices $\{S_{i_\alpha I_\beta}\}_{\alpha=1,\ldots,n;\beta=1,\ldots,m}$ as $\underline{\mathrm{S}}(\mathbf{i},\mathbf{I})$, and the $n \times m$ matrix of e–n expansion indices is $\underline{\mu}$.

We write $J_{n,m}$ as a sum of contributions from each group of n electrons and m nuclei in the system,

$$J_{n,m} = \frac{1}{n!m!}\sum_{\mathbf{i}}\sum_{\mathbf{I}} J_{n,m}(\mathbf{i},\mathbf{I}) = \sum_{\mathbf{i}}^{\text{s.v.}}\sum_{\mathbf{I}}^{\text{s.v.}} J_{n,m}(\mathbf{i},\mathbf{I}), \tag{5}$$

where summations with vector indices represent sums in which every component of the vector is a summation index, and "s.v." (for "sorted vector") indicates that the sum is restricted to vectors whose components are sorted, e.g., $i_1 < i_2 < \cdots < i_n$, which avoids redundant contributions [18]. The contribution from the n-electron and m-nucleus group $\{\mathbf{i},\mathbf{I}\}$ is

$$J_{n,m}(\mathbf{i},\mathbf{I}) = \sum_{\underline{\nu}}^{\text{u.t.}}\sum_{\underline{\mu}} \lambda^{\underline{\mathrm{P}}(\mathbf{i}),\underline{\mathrm{S}}(\mathbf{i},\mathbf{I})}_{\underline{\nu},\underline{\mu}} \prod^{\text{u.t.}} \underline{\Phi}^{\underline{\mathrm{P}}}_{\underline{\nu}}(\mathbf{i}) \prod \underline{\Theta}^{\underline{\mathrm{S}}}_{\underline{\mu}}(\mathbf{i},\mathbf{I}), \tag{6}$$

where λ are the linear parameters, summations with matrix indices represent sums in which every component of the matrix is a summation index $\prod$ acting on matrices implies the product of all of their components, and "u.t." means that the relevant operation is restricted to the upper-triangular portion of the e–e matrices involved, excluding the diagonal.

A. Symmetry properties of the linear parameters

Equation (5) imposes the condition that $J_{n,m}(\mathbf{i},\mathbf{I})$ must not depend on the specific ordering of the electrons and nuclei listed in $\mathbf{i}$ and $\mathbf{I}$. Let $\underline{\mathrm{A}}$ and $\underline{\mathrm{B}}$ be permutation matrices of respective sizes $n \times n$ and $m \times m$ such that $\underline{\mathrm{A}}\mathbf{i}$ and $\underline{\mathrm{B}}\mathbf{I}$ are integer vectors containing reordered electronic and nuclear indices. The value of $J_{n,m}(\underline{\mathrm{A}}\mathbf{i},\underline{\mathrm{B}}\mathbf{I})$ should therefore equal that of $J_{n,m}(\mathbf{i},\mathbf{I})$,

$$\begin{aligned} &J_{n,m}(\underline{\mathrm{A}}\mathbf{i},\underline{\mathrm{B}}\mathbf{I}) \\ &= \sum_{\underline{\nu}(\underline{\mathrm{A}}\mathbf{i})}^{\text{u.t.}}\sum_{\underline{\mu}(\underline{\mathrm{B}}\mathbf{I})} \lambda^{\underline{\mathrm{P}}(\underline{\mathrm{A}}\mathbf{i}),\underline{\mathrm{S}}(\underline{\mathrm{A}}\mathbf{i},\underline{\mathrm{B}}\mathbf{I})}_{\underline{\nu}(\underline{\mathrm{A}}\mathbf{i}),\underline{\mu}(\underline{\mathrm{A}}\mathbf{i},\underline{\mathrm{B}}\mathbf{I})} \prod^{\text{u.t.}} \underline{\Phi}^{\underline{\mathrm{P}}}_{\underline{\nu}}(\underline{\mathrm{A}}\mathbf{i}) \prod \underline{\Theta}^{\underline{\mathrm{S}}}_{\underline{\mu}}(\underline{\mathrm{A}}\mathbf{i},\underline{\mathrm{B}}\mathbf{I}) \\ &= \sum_{\underline{\mathrm{A}}\,\underline{\nu}\,\underline{\mathrm{A}}^{\mathrm{T}}}^{\text{u.t.}}\sum_{\underline{\mathrm{A}}\,\underline{\mu}\,\underline{\mathrm{B}}^{\mathrm{T}}} \lambda^{\underline{\mathrm{A}}\,\underline{\mathrm{P}}(\mathbf{i})\,\underline{\mathrm{A}}^{\mathrm{T}},\underline{\mathrm{A}}\,\underline{\mathrm{S}}(\mathbf{i},\mathbf{I})\,\underline{\mathrm{B}}^{\mathrm{T}}}_{\underline{\mathrm{A}}\,\underline{\nu}\,\underline{\mathrm{A}}^{\mathrm{T}},\underline{\mathrm{A}}\,\underline{\mu}\,\underline{\mathrm{B}}^{\mathrm{T}}} \prod^{\text{u.t.}} \underline{\mathrm{A}}\,\underline{\Phi}^{\underline{\mathrm{P}}}_{\underline{\nu}}(\mathbf{i})\,\underline{\mathrm{A}}^{\mathrm{T}} \prod \underline{\mathrm{A}}\,\underline{\Theta}^{\underline{\mathrm{S}}}_{\underline{\mu}}(\mathbf{i},\mathbf{I})\,\underline{\mathrm{B}}^{\mathrm{T}} \\ &= \sum_{\underline{\nu}}^{\text{u.t.}}\sum_{\underline{\mu}} \lambda^{\underline{\mathrm{A}}\,\underline{\mathrm{P}}(\mathbf{i})\,\underline{\mathrm{A}}^{\mathrm{T}},\underline{\mathrm{A}}\,\underline{\mathrm{S}}(\mathbf{i},\mathbf{I})\,\underline{\mathrm{B}}^{\mathrm{T}}}_{\underline{\mathrm{A}}\,\underline{\nu}\,\underline{\mathrm{A}}^{\mathrm{T}},\underline{\mathrm{A}}\,\underline{\mu}\,\underline{\mathrm{B}}^{\mathrm{T}}} \gamma\left[\underline{\Phi}^{\underline{\mathrm{P}}}_{\underline{\nu}}(\mathbf{i}),\underline{\mathrm{A}}\right] \prod^{\text{u.t.}} \underline{\Phi}^{\underline{\mathrm{P}}}_{\underline{\nu}}(\mathbf{i}) \prod \underline{\Theta}^{\underline{\mathrm{S}}}_{\underline{\mu}}(\mathbf{i},\mathbf{I}), \end{aligned} \tag{7}$$

where

$$\gamma\left[\underline{\Phi}^{\underline{\mathrm{P}}}_{\underline{\nu}}(\mathbf{i}),\underline{\mathrm{A}}\right] = \frac{\prod^{\text{u.t.}} \underline{\mathrm{A}}\,\underline{\Phi}^{\underline{\mathrm{P}}}_{\underline{\nu}}(\mathbf{i})\,\underline{\mathrm{A}}^{\mathrm{T}}}{\prod^{\text{u.t.}} \underline{\Phi}^{\underline{\mathrm{P}}}_{\underline{\nu}}(\mathbf{i})}, \tag{8}$$

which is $+1$ for basis sets consisting only of symmetric functions, while in the presence of antisymmetric basis functions it may be $+1$ or -1 depending on the precise permutation performed by $\underline{\mathrm{A}}$. Equating the right-hand sides of Eqs. (6) and (7) one finds that

$$\lambda^{\underline{\mathrm{P}}(\mathbf{i}),\underline{\mathrm{S}}(\mathbf{i},\mathbf{I})}_{\underline{\nu},\underline{\mu}} = \gamma\left[\underline{\Phi}^{\underline{\mathrm{P}}}_{\underline{\nu}}(\mathbf{i}),\underline{\mathrm{A}}\right] \lambda^{\underline{\mathrm{A}}\,\underline{\mathrm{P}}(\mathbf{i})\,\underline{\mathrm{A}}^{\mathrm{T}},\underline{\mathrm{A}}\,\underline{\mathrm{S}}(\mathbf{i},\mathbf{I})\,\underline{\mathrm{B}}^{\mathrm{T}}}_{\underline{\mathrm{A}}\,\underline{\nu}\,\underline{\mathrm{A}}^{\mathrm{T}},\underline{\mathrm{A}}\,\underline{\mu}\,\underline{\mathrm{B}}^{\mathrm{T}}}. \tag{9}$$

This equation represents the basic symmetry property of the linear parameters of the Jastrow factor, which implies that a parameter with a given set of superindices $\{\underline{\mathrm{P}}(\mathbf{i}),\underline{\mathrm{S}}(\mathbf{i},\mathbf{I})\}$ is determined by another parameter with a permuted set of superindices $\{\underline{\mathrm{A}}\,\underline{\mathrm{P}}(\mathbf{i})\,\underline{\mathrm{A}}^{\mathrm{T}},\underline{\mathrm{A}}\,\underline{\mathrm{S}}(\mathbf{i},\mathbf{I})\,\underline{\mathrm{B}}^{\mathrm{T}}\}$. This redundancy is removed by considering only one of the possible permutations of $\{\underline{\mathrm{P}}(\mathbf{i}),\underline{\mathrm{S}}(\mathbf{i},\mathbf{I})\}$. We call this particular permutation of $\{\underline{\mathrm{P}}(\mathbf{i}),\underline{\mathrm{S}}(\mathbf{i},\mathbf{I})\}$ the signature of the group of particles $\{\mathbf{i},\mathbf{I}\}$,

$$\{\underline{\tilde{\mathrm{P}}}(\mathbf{i}),\underline{\tilde{\mathrm{S}}}(\mathbf{i},\mathbf{I})\} = \{\underline{\mathrm{U}}\,\underline{\mathrm{P}}(\mathbf{i})\,\underline{\mathrm{U}}^{\mathrm{T}},\underline{\mathrm{U}}\,\underline{\mathrm{S}}(\mathbf{i},\mathbf{I})\,\underline{\mathrm{R}}^{\mathrm{T}}\}, \tag{10}$$

where the permutation matrices $\{\underline{\mathrm{U}},\underline{\mathrm{R}}\}$ are computed by applying a matrix-sorting algorithm [19] to $\{\underline{\mathrm{P}}(\mathbf{i}),\underline{\mathrm{S}}(\mathbf{i},\mathbf{I})\}$. In our terminology, the set of linear parameters whose superindices reduce to the same signature constitute a parameter channel. Only those parameters whose superindices equal the signature of a channel need be stored, and any other linear parameters in the channel can be computed from them via Eq. (9).

The signature $\{\underline{\tilde{\mathrm{P}}}(\mathbf{i}),\underline{\tilde{\mathrm{S}}}(\mathbf{i},\mathbf{I})\}$ may contain repeated entries such that there exist permutation matrices $\{\underline{\mathrm{A}},\underline{\mathrm{B}}\}$ that leave the signature unchanged,

$$\{\underline{\tilde{\mathrm{P}}}(\mathbf{i}),\underline{\tilde{\mathrm{S}}}(\mathbf{i},\mathbf{I})\} = \{\underline{\mathrm{A}}\,\underline{\tilde{\mathrm{P}}}(\mathbf{i})\,\underline{\mathrm{A}}^{\mathrm{T}},\underline{\mathrm{A}}\,\underline{\tilde{\mathrm{S}}}(\mathbf{i},\mathbf{I})\,\underline{\mathrm{B}}^{\mathrm{T}}\}, \tag{11}$$

in which case Eq. (9) becomes

$$\lambda^{\underline{\tilde{\mathrm{P}}}(\mathbf{i}),\underline{\tilde{\mathrm{S}}}(\mathbf{i},\mathbf{I})}_{\underline{\nu},\underline{\mu}} = \gamma\left[\underline{\Phi}^{\underline{\tilde{\mathrm{P}}}}_{\underline{\nu}}(\mathbf{i}),\underline{\mathrm{A}}\right] \lambda^{\underline{\tilde{\mathrm{P}}}(\mathbf{i}),\underline{\tilde{\mathrm{S}}}(\mathbf{i},\mathbf{I})}_{\underline{\mathrm{A}}\,\underline{\nu}\,\underline{\mathrm{A}}^{\mathrm{T}},\underline{\mathrm{A}}\,\underline{\mu}\,\underline{\mathrm{B}}^{\mathrm{T}}}. \tag{12}$$

Equation (12) is the symmetry constraint that relates linear parameters within a channel, which can be imposed as detailed in Sec. II C.

B. Indexing of basis functions

The components of $\underline{\nu}$ and $\underline{\mu}$ are the e–e and e–n expansion indices. We define the expansion indices so that they can each take any value between 1 and the e–e expansion order p, and between 1 and the e–n expansion order q, respectively. We factorize an optional cutoff function into Φ^P_ν and Θ^S_μ, so that

$$\Phi^P_\nu(\mathbf{r}) = f^P(\mathbf{r})\phi^P_\nu(\mathbf{r}), \tag{13}$$

for $\nu > 0$, and

$$\Theta^S_\mu(\mathbf{r}) = g^S(\mathbf{r})\theta^S_\mu(\mathbf{r}), \tag{14}$$

for $\mu > 0$, where f^P and g^S are the e–e and e–n cutoff functions and ϕ^P_ν and θ^S_μ are functions from a suitable basis set. This factorization allows an efficient implementation of localized Jastrow factor terms.

Additionally, we allow expansion indices to take a value of zero with the special meaning that $\Phi^P_0(\mathbf{r}) = \Theta^S_0(\mathbf{r}) = 1$ for all P, S, and $\mathbf{r}$. Note that these 0th functions do not contain cutoff functions. This allows us to construct terms with specialized functional forms, such as those involving dot products of vectorial quantities.

C. Constraints

Constraints on the parameters can be expressed in the form of a system of equations involving the linear parameters and the basis function parameters. We restrict our analysis to linear constraints on the linear parameters, and constraints that can be imposed on the nonlinear parameters contained in a basis function independently from the linear parameters and from nonlinear parameters in other basis functions.

Linear constraints on the linear parameters can be imposed using Gaussian elimination, as described in Ref. [7]. The matrix of coefficients may depend on the nonlinear parameters in the basis functions, if present, and the linear system is usually underdetermined, resulting in a subset of the parameters being determined by the values of the remaining parameters, which can be optimized directly.

When a constraint results in setting specific linear parameters to zero, it is more convenient simply to remove them from the list of linear parameters. This is accomplished by disallowing the indices $\underline{\nu}$ and $\underline{\mu}$ from taking the values corresponding to linear parameters. We call this an indexing constraint.

1. Symmetry and antisymmetry constraints

Symmetry constraints must always be imposed, otherwise the trial wave function is unphysical and calculations give erroneous results. Symmetry constraints amount to equalities between pairs of parameters as per Eq. (12). When two of these equalities relate the same pair of parameters with opposite signs, e.g., $\lambda_1 = \lambda_2$ and $\lambda_1 = -\lambda_2$, which implies $\lambda_1 = \lambda_2 = 0$, both parameters are eliminated using indexing constraints.

2. Constraints at e–e and e–n coalescence points

The Coulomb potential energy diverges when the positions of two electrons or an electron and a nucleus coincide. However, the local energy of an eigenstate of the Hamiltonian, including the exact ground-state wave function, is finite and constant throughout configuration space. Divergences in the local energy are therefore not a feature of the exact wave function and can lead to poor statistics in QMC calculations; hence it is important to avoid them. The kinetic energy must diverge to cancel out the potential energy and keep the local energy finite, which is achieved by demanding that the wave function obeys the Kato cusp conditions [8]. For any two charged particles i and j in a two- or three-dimensional system, these are

$$\left(\frac{1}{\Psi}\frac{\partial\hat{\Psi}}{\partial r_{ij}}\right)_{r_{ij}\to 0} = \frac{2q_iq_j\mu_{ij}}{d\pm 1}, \tag{15}$$

where $\hat{\Psi}$ denotes the spherical average of Ψ, q represents charge, $\mu_{ij} = m_im_j/(m_i+m_j)$ is the reduced mass, m represents mass, d is the dimensionality, and the positive sign in the denominator is for indistinguishable particles and the negative sign is for distinguishable particles. Fixed nuclei are regarded as having an infinite mass. Divergent interactions other than the Coulomb potential would give rise to different expressions.

It is common practice to impose the e–n cusp conditions on $\Psi_{\rm S}$ and the e–e cusp conditions on the Jastrow factor. This is because typical forms of $\Psi_{\rm S}$ explicitly depend on the e–n distances but not on the e–e distances. Our implementation allows the option of applying both types of cusp conditions to the Jastrow factor, which gives flexibility in the choice of $\Psi_{\rm S}$ and its properties. In particular, we impose the cusp conditions on a single Jastrow factor term and constrain all other terms in the Jastrow factor so that their contribution to the local kinetic energy is finite at e–e and e–n coalescence points. For nondivergent interaction potentials, such as most pseudopotentials, we simply require that the kinetic energy remains finite at coalescence points. Our implementation is also capable of not applying any constraints at e–e and e–n coalescence points since this is advantageous in some cases [20,21].

Imposing that the kinetic energy be finite at coalescence points is nontrivial if the Jastrow factor contains anisotropic functions. Consider the exponent of a Jastrow factor J near a point where two particles coalesce, be it two electrons or an electron and a nucleus. The dependence of J on coordinates other than those of the coalescing particles should be smooth in the vicinity of the coalescence point, and therefore one should be able to write $J = J(\mathbf{r})$, where $\mathbf{r}$ is the difference between the position vector of the two particles, and all remaining particles are held fixed.

The local kinetic energy is computed from two estimators, one involving $\nabla J(\mathbf{r})$ and the other $\nabla^2 J(\mathbf{r})$. We require both quantities to remain finite as $\mathbf{r}\to\mathbf{0}$. We expand the Jastrow factor in spherical harmonics, $J(\mathbf{r}) = \sum_{l=0}^{\infty}\sum_{m=-l}^{l} J^{(l,m)}$, with $J^{(l,m)} = f_{l,m}(r)Y_{l,m}(\theta,\phi)$, and the gradient and Laplacian of $J^{(l,m)}$ are

$$\begin{aligned}\nabla J^{(l,m)} &= f'_{l,m}(r)Y_{l,m}(\theta,\phi)\mathbf{u}_r + \frac{f_{l,m}(r)}{r}\frac{\partial Y_{l,m}(\theta,\phi)}{\partial\theta}\mathbf{u}_\theta \\ &\quad + \frac{f_{l,m}(r)}{r\sin\theta}\frac{\partial Y_{l,m}(\theta,\phi)}{\partial\phi}\mathbf{u}_\phi, \end{aligned}\tag{16}$$

$$\nabla^2 J^{(l,m)} = \left[f''_{l,m}(r) + \frac{2f'_{l,m}(r)}{r} - \frac{l(l+1)f_{l,m}(r)}{r^2}\right]Y_{l,m}(\theta,\phi). \tag{17}$$

Let us assume that $f_{l,m}(r)$ is finite at the origin and expand it to second order about $r=0$, $f_{l,m}(r) \approx f_{l,m}(0) + f'_{l,m}(0)r + f''_{l,m}(0)r^2/2$. Substituting into Eqs. (16) and (17), and ignoring contributions of $O(r)$ or higher, we arrive at

$$\begin{aligned}\nabla J^{(l,m)} &\approx f'_{l,m}(0)Y_{l,m}(\theta,\phi)\mathbf{u}_r \\ &\quad + \left[\frac{f_{l,m}(0)}{r} + f'_{l,m}(0)\right]\frac{\partial Y_{l,m}(\theta,\phi)}{\partial\theta}\mathbf{u}_\theta \\ &\quad + \left[\frac{f_{l,m}(0)}{r\sin\theta} + \frac{f'_{l,m}(0)}{\sin\theta}\right]\frac{\partial Y_{l,m}(\theta,\phi)}{\partial\phi}\mathbf{u}_\phi, \end{aligned}\tag{18}$$

$$\begin{aligned}\nabla^2 J^{(l,m)} &\approx \left\{\left[3 - \frac{l(l+1)}{2}\right]f''_{l,m}(0) + [2 - l(l+1)]\frac{f'_{l,m}(0)}{r}\right. \\ &\quad \left. - \, l(l+1)\frac{f_{l,m}(0)}{r^2}\right\}Y_{l,m}(\theta,\phi). \end{aligned}\tag{19}$$

The coefficient of the negative powers of r in Eqs. (18) and (19) must vanish for ∇J and $\nabla^2 J$ to be finite at the coalescence point. This gives rise to two conditions: (a) if $l \neq 0$ then $f_{l,m}(0) = 0$, and (b) if $l \neq 1$ then $f'_{l,m}(0) = 0$.

Application of the Kato cusp or finite kinetic energy constraints requires the construction of a linear system for each linear-parameter channel in each term of the Jastrow factor based on the above equations. Let $\mathcal{P}_{l,m}$ be an operator such that $\mathcal{P}_{l,m}\sum_{l=0}^{\infty}\sum_{m=-l}^{l} f_{l,m}(r)Y_{l,m}(\theta,\phi) = f_{l,m}(r)$. The cusp equations associated with the coalescence of electrons i and j have the form

$$\sum_{\underline{\nu}}^{\text{u.t.,o.c.}}\sum_{\underline{\mu}}^{\text{n.c.}}\lambda^{\underline{\text{P}},\underline{\text{S}}}_{\underline{\nu},\underline{\mu}}\left\{\frac{\partial\mathcal{P}_{0,0}\left[\Phi^{P_{ij}}_{\nu_{ij}}(\mathbf{r}_{ij})\right]}{\partial r_{ij}}\right\}_{r_{ij}=0} = \Gamma_{ij}, \qquad (20)$$

where Γ_{ij} is the right-hand side of Eq. (15). The label "o.c." (for "one contribution") denotes that the sum is restricted to values of the $\underline{\nu}$ sets such that the elements of the upper triangular portion of $\Phi^{\underline{\text{P}}}_{\underline{\nu}}(\mathbf{i})$ are all 1 except that corresponding to the electron pair formed by i and j, and the label "n.c." (for "no contribution") denotes that the sum is restricted to values of $\underline{\mu}$ such that all elements of $\Theta^{\underline{\text{S}}}_{\underline{\mu}}(\mathbf{i},\mathbf{I})$ are 1. These restrictions are trivially satisfied by e–e terms.

Parameters that do not contribute to Eq. (20) should be set by the condition that the kinetic energy does not diverge at e–e coalescence points, resulting in

$$\sum_{\underline{\nu}}^{\text{u.t.,e.p.}}\sum_{\underline{\mu}}^{\text{e.p.}}\lambda^{\underline{\text{P}},\underline{\text{S}}}_{\underline{\nu},\underline{\mu}}\left\{\mathcal{P}_{l,m}\left[\Phi^{P_{ij}}_{\nu_{ij}}(\mathbf{r}_{ij})\right]\right\}_{r_{ij}=0} = 0, \qquad (21)$$

for $l \neq 0$, and

$$\sum_{\underline{\nu}}^{\text{u.t.,e.p.}}\sum_{\underline{\mu}}^{\text{e.p.}}\lambda^{\underline{\text{P}},\underline{\text{S}}}_{\underline{\nu},\underline{\mu}}\left\{\frac{\partial\mathcal{P}_{l,m}\left[\Phi^{P_{ij}}_{\nu_{ij}}(\mathbf{r}_{ij})\right]}{\partial r_{ij}}\right\}_{r_{ij}=0} = 0, \qquad (22)$$

for $l \neq 1$. The anisotropy of $\Phi^{P_{ij}}_{\nu_{ij}}(\mathbf{r}_{ij})$ at $\mathbf{r}_{ij} = \mathbf{0}$ determines which of Eq. (21) and Eq. (22) need be imposed. The label e.p. (for "equal product") denotes that the sum is only over indices associated with electrons i and/or j, and these indices take only values such that the product of the pair of functions associated with ν_{ik} and ν_{jk} (μ_{iI} and μ_{jI} in the e–n case) is equal throughout the sum. Each set of distinct two-function products and each value of (l,m) correspond to different equations, and each set of possible values of the indices not summed over corresponds to a separate set of equations.

For the coalescence of electron i and nucleus I the cusp conditions take the form

$$\sum_{\underline{\nu}}^{\text{u.t.,n.c.}}\sum_{\underline{\mu}}^{\text{o.c.}}\lambda^{\underline{\text{P}},\underline{\text{S}}}_{\underline{\nu},\underline{\mu}}\left\{\frac{\partial\mathcal{P}_{l,m}\left[\Theta^{S_{iI}}_{\mu_{iI}}(\mathbf{r}_{iI})\right]}{\partial r_{iI}}\right\}_{r_{iI}=0} = \Gamma_{iI}, \qquad (23)$$

while the kinetic energy is kept finite if

$$\sum_{\underline{\nu},\underline{\mu}}^{\text{u.t.,e.p.}}\lambda^{\underline{\text{P}},\underline{\text{S}}}_{\underline{\nu},\underline{\mu}}\left\{\mathcal{P}_{l,m}\left[\Theta^{S_{iI}}_{\mu_{iI}}(\mathbf{r}_{iI})\right]\right\}_{r_{iI}=0} = 0 \qquad (24)$$

for $l \neq 0$ and

$$\sum_{\underline{\nu},\underline{\mu}}^{\text{u.t.,e.p.}}\lambda^{\underline{\text{P}},\underline{\text{S}}}_{\underline{\nu},\underline{\mu}}\left\{\frac{\partial\mathcal{P}_{l,m}\left[\Theta^{S_{iI}}_{\mu_{iI}}(\mathbf{r}_{iI})\right]}{\partial r_{iI}}\right\}_{r_{iI}=0} = 0 \qquad (25)$$

for $l \neq 1$. The equal-product constraint on the sum is now such that the sum is only over e–e indices associated with electron i and e–n indices associated with nucleus I, and these indices take values only such that the product of the pair of functions associated with ν_{ik} and μ_{kI} is equal throughout the sum. Again, each set of distinct two-function products corresponds to a different equation, and each set of possible values of the indices not being summed over corresponds to a separate set of equations.

Note that the equal-product constraints in the sums of Eqs. (20) and (23) reduce to the equal-product constraints described for the e–e–n f term of the Drummond-Towler-Needs (DTN) Jastrow factor in the appendix of Ref. [7] when natural-power basis functions are chosen.

3. Other constraints

It is possible to construct terms containing dot products by using appropriate constraints. For example, consider the basis functions $\Theta_1(\mathbf{r}) = x$, $\Theta_2(\mathbf{r}) = y$, and $\Theta_3(\mathbf{r}) = z$. In an e–n–n term we can restrict the indices so that $\underline{\mu}$ takes only the values (11), (22), (33), so that the contribution of electron i and nuclei I and J is $\mathbf{r}_{iI}\cdot\mathbf{r}_{iJ}$, provided we also apply a linear constraint that equates the three nonzero linear coefficients. Section III B gives a practical example of a term containing dot products, which is used in Sec. IV C4.

It is also possible to introduce Boys-Handy-style indexing, [22] where the sum of all e–e and e–n indices is restricted to be less than or equal to some fixed integer l. This is accomplished by setting the e–e and e–n expansion orders to l and then eliminating the parameters that violate the conditions via indexing restrictions.

III. BASIS FUNCTIONS AND TERMS

A. Basis sets and cutoff functions

Possibly the simplest basis set is the natural powers,

$$N_\nu(\mathbf{r}) = r^{\nu-1}, \qquad (26)$$

as used in the DTN Jastrow factor for the localized u, χ, and f terms [7]. These functions need to be cut off at some radius L, for which purpose the DTN Jastrow factor uses the polynomial cutoff function

$$D(\mathbf{r}) = (r-L)^C\Theta(L-r), \qquad (27)$$

where L is an optimizable parameter, C is a positive integer, and $\Theta(r)$ is the Heaviside step function. We also use a slightly different version of this cutoff function,

$$P(\mathbf{r}) = (1-r/L)^C\Theta(L-r), \qquad (28)$$

which should be numerically superior to $D(\mathbf{r})$.

A particular variant of $P(\mathbf{r})$ is the anisotropic cutoff function

$$A(\mathbf{r}) = (1-r/L)^C\Theta(L-r)\sum_i c_i\prod_\beta^d\left[\frac{\mathbf{r}\cdot\hat{\mathbf{u}}_\beta}{r}\right]^{p^{(i)}_\beta}, \qquad (29)$$

where L is an optimizable parameter, C is a positive integer, d is the dimensionality of the system, $\{\hat{\mathbf{u}}_\beta\}_{\beta=1,\dots,d}$ are unit vectors along d orthogonal directions, $\{c_i\}$ are real-valued constants, and $p^{(i)}_\beta$ are integer exponents, which are constrained so that $\sum_\beta^d p^{(i)}_\beta$ is the same for all values of i. This cutoff function is simply the product of an isotropic cutoff function and a spherical harmonic. For example, with $c_1 = 3$, $c_2 = -1$, $\mathbf{p}^{(1)} = (211)$, and $\mathbf{p}^{(2)} = (031)$, and the vectors pointing along the

Cartesian axes, we obtain

$$A(\mathbf{r}) = (1 - r/L)^C \Theta(L - r) \left[\frac{(3x^2 - y^2)yz}{r^4} \right], \tag{30}$$

which is proportional to a real spherical harmonic with $l = 4$. The advantage of describing anisotropy in the cutoff function rather than in the basis functions is that the common spherical harmonic can be factorized out of the sum over expansion indices, which reduces the computational cost. We allow different orientations to be used for different e–e or e–n dependency indices, which is useful to adapt the functional form to, e.g., the geometry of a molecule.

An alternative to the natural-power basis in finite systems is a basis of powers of fractions which tend to a constant as $r \to \infty$ and therefore do not need to be cut off. We define the basis

$$F_\nu(\mathbf{r}) = \left(\frac{r}{r^b + a} \right)^{\nu - 1}, \tag{31}$$

where a and b are real-valued optimizable parameters. Similar basis sets with $b = 1$ have been used in the literature, often in conjunction with Boys-Handy-style indexing [20,22–24], and this basis was used in Ref. [25] with an early implementation of the Jastrow factor presented here.

In extended systems it is important to use a basis that is consistent with the geometry of the simulation cell and has the periodicity of the system, such as a cosine basis,

$$C_\nu(\mathbf{r}) = \sum_{\mathbf{G} \in \nu\text{-th star}} \cos(\mathbf{G} \cdot \mathbf{r}), \tag{32}$$

where the $\mathbf{G}$ vectors are arranged in stars defined by the cell geometry. This basis is used in the DTN Jastrow factor for the extended p and q terms.

A suitable basis set for building specialized terms containing dot products is

$$V_\nu(\mathbf{r}) = r^{\mathrm{INT}[(\nu-1)/d]} \frac{\mathbf{r} \cdot \hat{\mathbf{u}}_{\mathrm{MOD}(\nu-1,d)+1}}{r}, \tag{33}$$

where d is the dimensionality of the system and $\{\hat{\mathbf{u}}_\beta\}_{\beta=1,\ldots,d}$ are unit vectors parallel to the d Cartesian axes. A term constructed using these functions with appropriate index-restriction constraints would consist of dot products between two vectors multiplied by a natural-power expansion in their moduli.

B. Terms and notation

We employ a condensed notation to refer to Jastrow terms that use certain basis functions, cutoff functions, and constraints. Each term is represented by a single capital letter, with n and m as subindices. Any other relevant information is given as a superindex. Typically we use expansion orders p and q of 7–9 for two-body terms, 4–5 for three-body terms, and 2–3 for four-body terms, except when indicated otherwise.

For simple Jastrow terms we use the natural power basis functions N_ν and the polynomial cutoff functions P or D. We refer to these terms as $N_{n,m}$. $N_{2,0}$, $N_{1,1}$, and $N_{2,1}$ are the equivalent of the DTN u, χ, and f terms, respectively. In the $N_{2,1}$ term, and in any term where more than one electron and one or more nuclei are involved, we choose not to apply e–e cutoff functions, relying instead on the e–n cutoffs to fulfill this role. Additional $N_{n,m}$ terms used in this paper that were not part of the DTN Jastrow factor are $N_{1,2}$, $N_{3,0}$, $N_{1,3}$, $N_{2,2}$, $N_{3,1}$, and $N_{4,0}$. In $N_{n,m}$ we typically use a truncation order in the cutoff function of $C = 3$.

We use $A^{\text{s.h.}}_{n,m}$ to refer to the anisotropic variant of $N_{n,m}$. The $A^{\text{s.h.}}_{n,m}$ term consists of natural power basis functions N_ν and the anisotropic cutoff function A, and "s.h." is a placeholder for describing the spherical harmonic. For example, for the highly anisotropic N_2 molecule we use terms such as $A^z_{1,1}$, $A^{z^2}_{1,1}$, $A^z_{2,1}$, and $A^{z^2}_{2,1}$.

In finite systems we also use the F_ν basis functions in terms without explicit cutoff functions which we call $F_{n,m}$, or $F^{b=1}_{n,m}$ when we force $b = 1$ in the basis functions. In some systems it is useful to apply Boys-Handy-style indexing to $F^{b=1}_{n,m}$, and we refer to the resulting term as $B_{n,m}$.

In extended systems we make use of the cosine basis functions C_ν in terms denoted $C_{n,m}$, where we choose expansion orders so that at least as many $\mathbf{G}$ vectors as electrons in each spin channel are included in the expansion.

To test the flexibility of our implementation we have designed an e–e–n–n Jastrow term for describing the correlations associated with van der Waals interactions, which we call $V_{2,2}$. This term is capable of distinguishing between configurations where the electron-nucleus relative position vectors $\mathbf{r}_{iI}$ and $\mathbf{r}_{jJ}$ are parallel from those where they are antiparallel. Introducing a dot product achieves this effect, and $V_{2,2}$ has the following functional form:

$$V_{2,2} = \frac{1}{2} \sum_{i \neq j}^{N} \sum_{I \neq J}^{M} P(r_{iI}) P(r_{jJ}) \sum_{\nu_{ij}}^{p} \sum_{\mu_{iI},\mu_{jJ}}^{q} \lambda_{\nu_{ij}\mu_{iI}\mu_{iJ}} \times N_{\nu_{ij}}(r_{ij}) N_{\mu_{iI}}(r_{iI}) N_{\mu_{jJ}}(r_{jJ}) \mathbf{r}_{iI} \cdot \mathbf{r}_{jJ}. \tag{34}$$

We require basis functions to be scalars in our Jastrow factor, so the dot product is separated into its components. Hence, we construct the $V_{2,2}$ term using V_ν for the e–n basis with P as the e–n cutoff functions, and N_ν for the e–e basis. We allow e–n indices to be zero and apply a number of constraints on the linear parameters: (a) we eliminate all index sets except those in which the e–n indices are of the form $\underline{\mu} = \binom{k\ 0}{0\ l}$ or $\binom{0\ k}{l\ 0}$, with $k,l > 0$; (b) we eliminate all parameters that do not satisfy $\mathrm{MOD}(k-1,3) = \mathrm{MOD}(l-1,3)$; (c) we equate any two linear parameters $\lambda_{\nu,\underline{\mu}_1}$ and $\lambda_{\nu,\underline{\mu}_2}$ if the zeros of $\underline{\mu}_1$ and $\underline{\mu}_2$ are in the same position and their nonzero components satisfy $\mathrm{INT}[(k_1 - 1)/3)] = \mathrm{INT}[(k_2 - 1)/3]$ and $\mathrm{INT}[(l_1 - 1)/3)] = \mathrm{INT}[(l_2 - 1)/3]$. These constraints are applied in addition to the generic symmetry and cusp constraints. Table I summarizes the notation for the terms we have introduced.

IV. RESULTS

In the present work we have used a variety of methods to optimize our Jastrow factors, namely, variance minimization, minimization of the mean absolute deviation of the local energy with respect to the median energy, and linear least-squares energy minimization [26,27]. All of our final wave functions are energy-minimized except where otherwise stated. Starting with the Hartree-Fock (HF) wave function, we progressively introduce Jastrow terms and reoptimize all of the parameters simultaneously. Optimizing the Jastrow factor term by term is unnecessary in practical applications, but here it allows us

TABLE I. Notation for Jastrow terms correlating n electrons and m nuclei using different basis functions.

Name	Basis set	Cutoff function	Special constraints
$N_{n,m}$	Natural powers	Polynomial	None
$F_{n,m}$	Powers of $r/(r^b+a)$	None	None
$B_{n,m}$	Powers of $r/(r+a)$	None	Boys-Handy-style indexing
$A^{\text{s.h.}}_{n,m}$	Natural powers	Anisotropic polynomial	None
$C_{n,m}$	Cosines	None	None
$V_{n,m}$	Natural powers times unit vectors	Polynomial	Dot product

to understand the importance of the different terms. We refer to the total number of optimizable parameters in the wave function as N_p.

The correlation energy is defined as the difference between the HF energy and the exact energy, $E_{\text{HF}} - E_{\text{exact}}$. The fraction of the correlation energy retrieved in a VMC calculation with a given trial wave function Ψ,

$$f_{\text{CE}}[\Psi] = \frac{E_{\text{HF}} - E_{\text{VMC}}[\Psi]}{E_{\text{HF}} - E_{\text{exact}}}, \tag{35}$$

is a measure of the quality of Ψ. We refer to the difference between the DMC and HF energies as the DMC correlation energy, $E_{\text{HF}} - E_{\text{DMC}}[\Psi]$. The fraction of the DMC correlation energy retrieved in VMC,

$$f_{\text{DCE}}[\Psi] = \frac{E_{\text{HF}} - E_{\text{VMC}}[\Psi]}{E_{\text{HF}} - E_{\text{DMC}}[\Psi]}, \tag{36}$$

measures the quality of the Jastrow factor, since a perfect Jastrow factor would make the VMC and DMC energies coincide [28]. We define the fraction of the remaining DMC correlation energy recovered by a wave function Ψ_2 with respect to another Ψ_1 as

$$\frac{E_{\text{VMC}}[\Psi_1] - E_{\text{VMC}}[\Psi_2]}{E_{\text{VMC}}[\Psi_1] - E_{\text{DMC}}[\Psi_2]}. \tag{37}$$

The local energy of an electronic configuration $\mathbf{R}$ is defined as $E_L(\mathbf{R}) = \Psi_{\text{T}}^{-1}(\mathbf{R})\hat{H}(\mathbf{R})\Psi_{\text{T}}(\mathbf{R})$, where $\hat{H}(\mathbf{R})$ is the Hamiltonian operator. The variance of the local energies encountered in a VMC calculation, which we shall refer to as the VMC variance, tends to its lower bound of zero as Ψ_{T} tends to an eigenstate of the Hamiltonian and is thus a measure of the overall quality of the trial wave function.

A. Homogeneous electron gases

1. *One-dimensional homogeneous electron gas*

We have studied a 1D HEG of density parameter $r_s = 5$ a.u. consisting of 19 electrons subject to periodic boundary conditions using a single Slater determinant of plane-wave orbitals. The ground-state energy of an infinitely thin 1D HEG in which electrons interact by the full Coulomb potential is independent of the magnetic state, and hence we have chosen all the electrons to have the same spin. This system is unusual in that the nodal surface of the trial function is exact, and therefore DMC gives the exact ground-state energy, which we have estimated to be $-0.2040834(3)$ a.u. per electron. Excellent results were reported for this system in Refs. [29,30] using wave functions with e–e backflow transformations [6,13] which preserve the (exact) nodal surface of the Slater determinant.

We have investigated the improvement in VMC results when various terms are added to an e–e Jastrow factor $J = N_{2,0} + C_{2,0}$, both with and without backflow transformations. In the absence of backflow, we find that including $N_{3,0}$, $C_{3,0}$, or $N_{3,0} + C_{3,0}$ improves the VMC energy, while the subsequent addition of $C_{4,0}$ yields no further gain. VMC gives an almost exact energy with backflow and $J = N_{2,0}$, and therefore no further reduction is possible by including more Jastrow terms. However, the addition of $N_{3,0} + C_{3,0}$ reduces the VMC variance by a factor of five, giving a variance that is an order of magnitude smaller than that reported in Ref. [29] for a similar calculation.

2. *Two-dimensional homogeneous electron gas*

We have studied a paramagnetic 2D HEG with 42 electrons per simulation cell at $r_s = 35$ a.u., which lies close to the Wigner crystallization density predicted by Drummond and Needs [31]. Kwon *et al.* [13] found that three-electron correlations are important at low densities, and that the effect of a three-electron Jastrow factor on the VMC energy is comparable to that of backflow. This makes low densities appealing for testing higher-rank Jastrow terms.

The VMC energy and variance obtained using different Jastrow factors with and without backflow is plotted in Fig. 1.

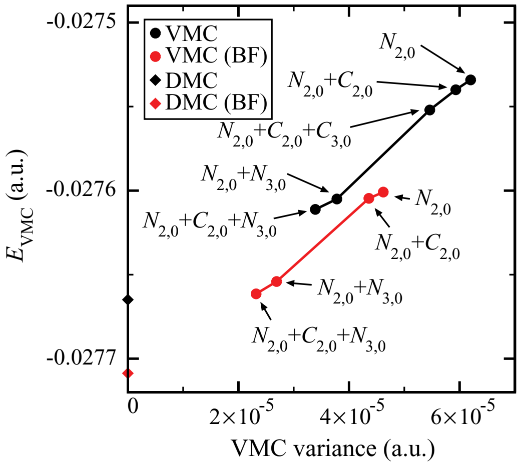

FIG. 1. (Color online) VMC energies E_{VMC} against the VMC variance for the 2D HEG at $r_s = 35$ a.u. using different Jastrow factors, along with the DMC energies for reference. The error bars are smaller than the size of the symbols, and "(BF)" indicates the use of backflow.

The addition of an $N_{3,0}$ term to $J = N_{2,0}$ recovers 81% of the remaining DMC correlation energy without backflow and 49% with backflow. The $C_{2,0}$ term further reduces both the VMC energy and variance. The use of a $C_{3,0}$ term recovers 10% of the remaining DMC correlation energy when added to $J = N_{2,0} + C_{2,0}$, but it was not used further since the lack of a cutoff function makes calculations with $C_{3,0}$ too costly for the little benefit it provides.

We have also performed DMC calculations using two different Jastrow factors in the presence of backflow in order to quantify the indirect effect of the quality of the Jastrow factor on the quality of the nodes of the wave function. We obtain a DMC energy of $-0.0277072(1)$ a.u. per electron for $J = N_{2,0}$, and a lower energy of $-0.0277087(1)$ a.u. per electron for $J = N_{2,0} + N_{3,0} + C_{2,0}$. This supports the idea that a better Jastrow factor allows the backflow transformation to shift its focus from the "bulk" of the wave function to its nodes, thus improving the DMC energy.

B. Be, B, and O atoms

While excellent descriptions of these atoms can be obtained within VMC and DMC using multideterminant wave functions with backflow correlations [25,32], we have used single-determinant wave functions since we are interested only in the effects of the Jastrow factor. We have studied the ground states of the Be, B, and O atoms, corresponding to ^{1}S, ^{2}P, and ^{3}P electronic configurations, respectively. The ATSP2K code [33] was used to generate numerical single-electron HF orbitals tabulated on a radial grid. We have investigated the use of Jastrow factors with up to four-body terms, but we have not used backflow for these systems. The energies of Chakravorty *et al.* [34] have been used as "exact" reference values.

We obtain lower single-determinant VMC energies for the Be, B, and O atoms with $J = F_{2,0} + F_{1,1} + F_{2,1}$ than reported in Refs. [32,35]. We obtain further small improvements in the VMC energies by including either $F_{3,0}$ or $F_{3,1}$ Jastrow terms, but their combination, $F_{3,0} + F_{3,1}$, is not found to be advantageous over using the terms individually. This indicates that $F_{3,0}$ and $F_{3,1}$, the latter of which provides a slightly lower VMC energy than the former, have nearly the same effect in atoms. These three-electron terms should be particularly useful in describing correlations involving electrons in the atomic core region. We expect $F_{3,1}$ to be more useful than $F_{3,0}$ in molecules and solids because it should be able to adapt to the different length scales in these systems, whereas $F_{3,0}$ offers a homogeneous description of three-electron correlations. We have investigated the effect of adding a $F_{4,1}$ term in Be and O, but it does not reduce the VMC energy or variance when added to $J = F_{2,0} + F_{1,1} + F_{2,1} + F_{3,1}$.

Our best VMC energies of $-14.6522(1)$ a.u. (Be), $-24.6309(2)$ a.u. (B), and $-75.0381(3)$ a.u. (O) correspond to fractions of the DMC correlation energy of 94.0(1)%, 91.8(1)%, and 94.6(1)%, respectively. The VMC energies are compared with the best single-determinant nonbackflow VMC values we could find in the literature in Table III.

C. BeH, N_2, H_2O, and H_2 molecules

The BeH, N_2, H_2O, and H_2 molecules are strongly inhomogeneous and anisotropic systems. We have used basis sets of moderate quality for the single-electron orbitals of BeH and N_2 in order to investigate the extent to which the Jastrow factor can compensate for the deficiencies of the basis sets, especially via one-electron terms $N_{1,m}$. For H_2O and H_2 we have used very good basis sets. We have also tested anisotropic Jastrow factors in N_2, and a van der Waals-like Jastrow factor for H_2.

1. BeH molecule

We have studied the all-electron BeH molecule in the $^2\Sigma^+$ ground state configuration at a bond length of 2.535 a.u. [36]. We have used a single-determinant wave function containing Slater-type orbitals generated with the ADF package [37], with which we obtain a reference DMC energy of $-15.24603(4)$ a.u.

The addition of $N_{1,2}$ to $J = N_{2,0} + N_{1,1} + N_{2,1}$ recovers 11% of the remaining DMC correlation energy. We find no significant gain from adding either an $N_{2,2}$ term or an $N_{3,1}$ term to $J = N_{2,0} + N_{1,1} + N_{2,1} + N_{1,2}$.

2. N_2 molecule

We have studied the $^1\Sigma_g^+$ ground state of the N_2 molecule at the experimental bond length of 2.074 a.u. [36] HF orbitals were generated in a Slater-type basis using the ADF package [37]. Our VMC results for different Jastrow factors are given in Table II along with relevant reference energies.

Adding an $N_{1,2}$ term to $J = N_{2,0} + N_{1,1} + N_{2,1}$ factor recovers 33% of the remaining DMC correlation energy and leads to a significant reduction in the VMC variance. The subsequent addition of $N_{2,2}$ provides a reduction in the VMC energy of 13% of the remaining DMC correlation energy. We have tested adding $N_{3,0}$, $N_{3,1}$, and $N_{4,0}$ terms to $J = N_{2,0} + N_{1,1} + N_{2,1} + N_{2,2}$, but neither of these yield any improvements in the VMC energy.

The anisotropy of this system is expected to be captured by terms containing e–n functions that treat the bond as a special direction. We have aligned the z axis of our reference frame along the N–N bond in our calculations, and $A^z_{1,1}$ is then the simplest explicitly anisotropic term that reflects the geometry of the system. The $A^x_{1,1}$ and $A^y_{1,1}$ terms must be zero by symmetry, and we have therefore not used them. There are five spherical harmonics with $l = 2$, which are respectively proportional to xy, xz, yz, $x^2 - y^2$, and $-x^2 - y^2 + 2z^2$. In our calculations we find that only the last one of these, which we refer to as z^2, has a significant effect on the VMC energy.

The VMC energy with $J = N_{2,0} + N_{1,1} + N_{2,1} + A^z_{1,1}$ is within statistical uncertainty of that with $J = N_{2,0} + N_{1,1} + N_{2,1} + N_{1,2}$, but the former Jastrow factor contains about a third fewer parameters than the latter. The combination of the $N_{1,2}$ and $A^z_{1,1}$ terms into $J = N_{2,0} + N_{1,1} + N_{2,1} + N_{1,2} + A^z_{1,1}$ does not improve the VMC energy compared with the other two Jastrow factors. These results suggest that the terms $N_{1,2}$ and $A^z_{1,1}$ play similar roles in the wave function, which we find reasonable since $N_{1,2}$, although constructed from isotropic basis functions, contains the right variables to capture the symmetry of the molecule in much the same way as $A^z_{1,1}$ does. We have plotted the $A^z_{1,1}$ term for $J = N_{2,0} + N_{1,1} + N_{2,1} + A^z_{1,1}$ and the $N_{1,2}$ term for $J = N_{2,0} + N_{1,1} + N_{2,1} + N_{1,2}$ in Fig. 2, where the similarity between the terms can be seen. The value of the $N_{1,2}$ term is roughly the same as that of $A^z_{1,1}$ offset

TABLE II. VMC energies (E) and variances (V) for the N_2 molecule using different Jastrow factors, including explicitly anisotropic terms. We have used a bond length of $r_{NN} = 2.074$a.u [36].

	N_p	E (a.u.)	V (a.u.)	f_{CE} (%)	f_{DCE} (%)
HF limit from Ref. [38]		– 108.9929		0	0
$N_{2,0}$	18	– 109.102(1)	– 5.275(4)	– 19.9(2)	– 21.3(2)
$N_{2,0} + N_{1,1}$	27	– 109.3739(6)	– 3.681(3)	– 69.4(1)	– 74.3(2)
$N_{2,0} + N_{1,1} + N_{1,2}$	49	– 109.3796(6)	– 3.595(2)	– 70.4(1)	– 75.4(2)
$N_{2,0} + N_{1,1} + N_{2,1}$	80	– 109.4441(4)	– 1.667(2)	– 82.16(7)	– 87.9(1)
$N_{2,0} + N_{1,1} + N_{2,1} + N_{1,2}$	102	– 109.4644(4)	– 1.149(2)	– 85.85(7)	– 91.9(1)
$N_{2,0} + N_{1,1} + N_{2,1} + N_{1,2} + N_{2,2}$	219	– 109.4697(4)	– 1.088(3)	– 86.82(7)	– 92.9(1)
$N_{2,0} + N_{1,1} + N_{2,1} + N_{1,2} + N_{2,2} + N_{3,0}$	260	– 109.4702(3)	– 1.083(2)	– 86.91(5)	– 93.0(1)
$N_{2,0} + N_{1,1} + A^{z}_{1,1}$	36	– 109.3770(6)	– 3.670(2)	– 69.9(1)	– 74.9(2)
$N_{2,0} + N_{1,1} + N_{2,1} + A^{z}_{1,1}$	89	– 109.4660(3)	– 1.116(2)	– 86.14(5)	– 92.2(1)
$N_{2,0} + N_{1,1} + N_{2,1} + A^{z}_{1,1} + A^{z^2}_{1,1}$	97	– 109.4669(3)	– 1.073(2)	– 86.31(5)	– 92.4(1)
$N_{2,0} + N_{1,1} + N_{2,1} + A^{z}_{1,1} + A^{z}_{2,1}$	142	– 109.4707(3)	– 1.072(2)	– 87.00(5)	– 93.1(1)
$N_{2,0} + N_{1,1} + N_{2,1} + A^{z}_{1,1} + A^{z^2}_{1,1} + A^{z}_{2,1} + A^{z^2}_{2,1}$	191	– 109.4714(3)	– 1.036(4)	– 87.13(5)	– 93.3(1)
SD VMC from Ref. [35][a]		– 109.4520(5)		– 83.59(9)	– 89.5(2)
SD DMC		– 109.5060(7)		– 93.4(1)	100.0(2)
Exact from Ref. [38]		– 109.5421		100	

[a] For $r_{NN} = 2.075$ a.u. We do not expect that this small difference in bond length will affect the comparison between energies significantly.

by a positive amount, and this shift is likely to be compensated for by the other Jastrow factor terms. Both terms increase the value of the wave function in the outer region of the molecule with respect to that in the bond region.

We have added different combinations of anisotropic terms to $J = N_{2,0} + N_{1,1} + N_{2,1}$. The e–e–n $A^{z}_{2,1}$ term retrieves less correlation energy than the e–n $A^{z}_{1,1}$ term. Combining terms with spherical harmonics of $l = 1$ and $l = 2$ improves the VMC energy significantly with respect to using $l = 1$ only. The anisotropic Jastrow factor $J = N_{2,0} + N_{1,1} + N_{2,1} + A^{z}_{1,1} + A^{z^2}_{1,1} + A^{z}_{2,1} + A^{z^2}_{2,1}$, which contains up to e–e–n correlations and has 191 optimizable parameters, recovers 93.3(1)% of the DMC correlation energy, which is greater than the 93.0(1)% retrieved by our best isotropic Jastrow factor $J = N_{2,0} + N_{1,1} + N_{2,1} + N_{1,2} + N_{2,2} + N_{3,0}$, which includes more costly e–e–n–n and e–e–e correlations and contains 260 optimizable parameters. We conclude that anisotropic functions are an important tool in the construction of compact Jastrow factors for strongly anisotropic systems.

Toulouse and Umrigar obtained 90% of the DMC correlation energy with a single-determinant wave function in Ref. [35], and with our best Jastrow factor we retrieve 93% of the DMC correlation energy. We have also optimized a single-determinant backflow wave function with our best Jastrow factor, and we obtain a VMC energy of $-109.4820(6)$ a.u. (89% of the correlation energy), which is of similar accuracy to the multideterminant VMC energy of $-109.4851(3)$ a.u. (89.6% of the correlation energy) obtained by Toulouse and Umrigar.

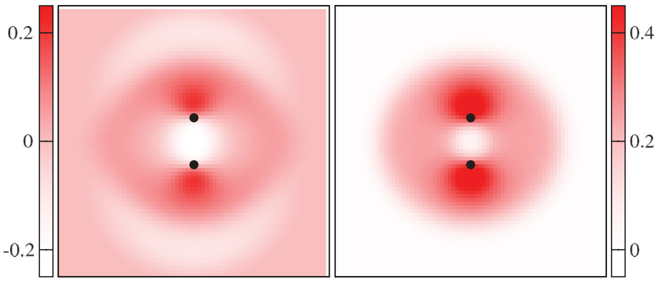

FIG. 2. (Color online) Plots of the $A^{z}_{1,1}$ term (left) and $N_{1,2}$ term (right) of different Jastrow factors for N_2 as a function of the position of an electron in a plane containing the nuclei (black circles).

3. H_2O molecule

Single-particle spin-unrestricted HF orbitals for the $^1A_1$ ground state of H_2O were generated using the CRYSTAL Gaussian basis set code [39]. The basis set for O contains 14 s, 9 p, and 4 d orbitals, and that for H contains 8 s, 4 p, and 3 d orbitals. Electron-nucleus cusps have been added using the scheme of Ma *et al.* [40]. We have simulated a water molecule with a bond length of $r_{OH} = 1.8088$ a.u. and a bond angle of $\angle_{HOH} = 104.52^\circ$ [41].

Adding an $N_{1,2}$ term to $J = N_{2,0} + N_{1,1} + N_{2,1}$ gives only a very small improvement for H_2O, compared with the more substantial improvements obtained with this e–n–n term for BeH and N_2. The $N_{1,2}$ term acts as a correction to the single-electron orbitals, and we believe that it is unimportant in this case because we have used very accurate HF orbitals, whereas the single-electron orbitals used for BeH and N_2 are considerably less accurate. We find additional small improvements to the energy of H_2O from adding $N_{3,0}$ and $N_{3,1}$ terms to $J = N_{2,0} + N_{1,1} + N_{2,1}$.

Clark *et al.* obtained 92% of the DMC correlation energy with a single-determinant wave function in Ref. [42], and with our best Jastrow factor we recover 95.5% of the DMC correlation energy.

4. H_2 molecule

The energy of the first triplet excited state ($^3\Sigma_u^+$) of H_2 has a very shallow minimum corresponding to a large bond length of nearly 8 a.u. Although the exchange interaction falls exponentially with increasing internuclear separation, Kolos and Wolniewicz found that it contributed significantly to the energy even at the large distance of 10 a.u. [43]. The strong interplay between the attractive dispersion forces and the repulsive exchange interaction requires that both be accounted for to afford an accurate description of the triplet state. This makes the system appealing for studying the construction of four-body Jastrow factor terms to describe van der Waals-like interactions.

We generated numerical HF orbitals tabulated on an elliptical grid using the 2DHF package [44]. HF theory predicts no binding for the triplet state at any separation, and therefore any binding that occurs in VMC can be attributed to the Jastrow factor. We have studied the H_2 molecule in its triplet spin state at the internuclear distance of 7.8358 a.u. This separation and the corresponding energy of -1.0000208957 a.u. were found by fitting a quadratic function to the data of Staszewska and Wolniewicz [45].

Previous QMC calculations on H_2 at different interatomic distances have used Jastrow factors with up to four-body correlations where the cusp conditions were not enforced [14,20], instead relying on the variance minimization method to find parameter values that approximately satisfy the cusp conditions. This was found to be advantageous for this system because the additional variational freedom yielded a better description in VMC than when the cusp conditions were obeyed exactly [21]. The violation of the cusp conditions is potentially catastrophic in DMC calculations, but previous studies have restricted the use of such terms to VMC.

For H_2 we have optimized Jastrow factors consisting of the single e–e–n–n terms $V_{2,2}$, $F_{2,2}^{b=1}$, and $B_{2,2}$ (see Table I) at several expansion orders, where no constraints are enforced at e–e or e–n coalescence points. We have used variance minimization for these Jastrow factors as we found that it produces better results than energy minimization. We have also optimized Jastrow factors consisting of different sums of terms which satisfy the cusp conditions using energy minimization. The results are shown in Fig. 3 .

We have performed the DMC calculations using our best $B_{2,2}$ Jastrow factor and obtain a reference DMC energy of $-1.0000207(1)$ a.u. We have not encountered any statistical problems in the DMC calculations with this cusp-violating wave function. Such issues can occur when the local energy presents a negative divergence in a region of configuration space with a significant probability of being sampled. We have verified that our wave function causes a negative divergence in the local energy when an electron coalesces with a nucleus, a point at which the wave function is likely to be relatively large. We therefore conclude that the region of influence of this divergence is sufficiently small that statistical problems do not arise in practice.

The $F_{2,2}^{b=1}$ and $B_{2,2}$ terms only differ in that the latter uses Boys-Handy-style indexing, which yields slightly lower VMC energies than standard indexing in most cases for a fixed number of parameters. Our best $F_{2,2}^{b=1}$ and $B_{2,2}$ Jastrow factors retrieve 99% of the DMC correlation energy in VMC.

The $V_{2,2}$ term is designed to describe van der Waals correlations, and contains e–e functions which introduce other correlations. Our best $V_{2,2}$ term recovers 92% of the DMC correlation energy, offering a good description of the system without reaching the accuracy of the more generic $F_{2,2}^{b=1}$ and $B_{2,2}$ terms.

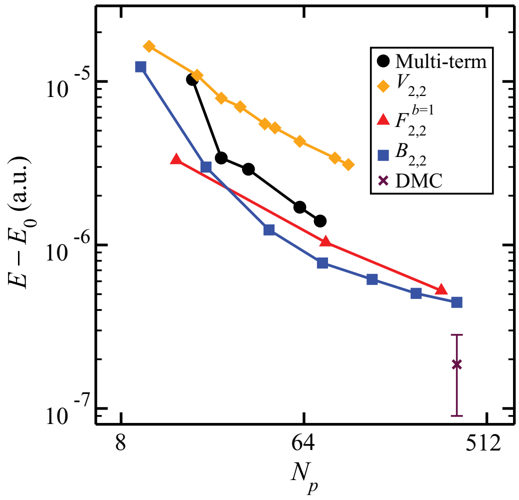

FIG. 3. (Color online) Difference between the VMC and exact energy against the number of wave function parameters for the H_2 triplet ground state using different Jastrow factors. Only the “multiterm” Jastrow factor enforces the cusp conditions. The error bars are smaller than the size of the symbol where not shown. All of the energies in this plot are lower than -1 a.u., and therefore the wave functions predict binding.

A $V_{2,2}$ term without e–e functions consists of contributions proportional to $\mathbf{r}_{iI} \cdot \mathbf{r}_{jJ}$, where the prefactors depend explicitly on r_{iI} and r_{jJ}, and implicitly on r_{IJ}, and this functional form is that of a dipole-dipole interaction potential. Our best such $V_{2,2}$ term retrieves 69% of the DMC correlation energy, which amounts to 0.0000262(3) a.u., and we regard this as a measure of the pure van der Waals correlation energy of this system.

The multiterm Jastrow factors contain the usual $N_{2,0}$, $N_{1,1}$, $N_{1,2}$, and $N_{2,1}$ terms, and for each combination of these we have added a $V_{2,2}$ term without e–e functions obeying the cusp conditions to study its effect. $J = N_{2,0}$ retrieves 44%

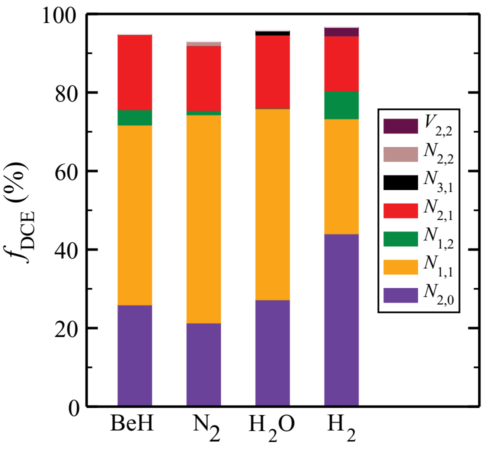

FIG. 4. (Color online) Fraction of the DMC correlation energy retrieved by different Jastrow factor terms, added in the specific order depicted in the graph (starting from the bottom), for the molecules studied in this paper.

TABLE III. Best single-determinant nonbackflow VMC energies (a.u.) found in the literature and those from this work, along with single-determinant DMC and exact energies for reference.

System	This work	Literature	DMC	Exact
1D HEG ($r_s = 5$, $N = 19$)	– 0.2040833(2)		– 0.2040834(3)	– 0.2040834(3)
2D HEG ($r_s = 35$, $N = 42$)	– 0.0276112(6)		– 0.0277087(1)	
Be	– 14.6522(1)	– 14.64972(5)[a]	– 14.65717(4)	– 14.66736
B	– 24.6309(2)	– 24.62936(5)[a]	– 24.64002(6)	– 24.65391
O	– 75.0381(3)	– 75.0352(1)[a]	– 75.0511(1)	– 75.0673
BeH	– 15.2412(3)	– 15.228(1)[b]	– 15.24603(4)	– 15.2482
N_2	– 109.4714(3)	– 109.4520(5)[c]	– 109.5060(7)	– 109.5421
H_2O	– 76.4068(2)	– 76.3938(4)[d]	– 76.4226(1)	– 76.438
H_2 ($^3\Sigma_u^+$)	– 1.00002045(3)		– 1.0000207(1)	– 1.0000208957

[a] Reference [35].

[b] Reference [46].

[c] Reference [35] (using a slightly different bond length).

[d] Reference [42].

of the DMC correlation energy, and adding the $V_{2,2}$ term retrieves 85% of the remaining DMC correlation energy. The effectiveness of $V_{2,2}$ progressively drops as more terms are added, and it retrieves 43% of the remaining DMC correlation energy when added to $J = N_{2,0} + N_{1,1} + N_{2,1} + N_{1,2}$. In all cases, $V_{2,2}$ is found to lower the VMC energy by a larger amount than any of the $N_{n,m}$ terms.

Our best multiterm cusp-enforcing Jastrow factor retrieves 97% of the DMC correlation energy with 77 wave-function parameters, comparable with the 98% retrieved with the cusp-violating $F_{2,2}^{b=1}$ and $B_{2,2}$ terms with a similar number of parameters. For larger systems where van der Waals interactions are important, we expect the violation of cusp conditions to cause statistical problems, and the $V_{2,2}$ term would become an effective way of improving the description of the system in a multiterm Jastrow factor.

D. Discussion of molecular results

In Fig. 4 we have plotted the fraction of the DMC correlation energy retrieved by different Jastrow factor terms for BeH, N_2, H_2O, and H_2. Our purpose is to visualize the importance of different terms in different systems, and to this end we do not include anisotropic or cusp-violating terms, and we consider only the addition of terms in a specific order.

The $N_{2,0}$ term represents the simplest description of electronic correlations and typically retrieves 20%–25% of the DMC correlation energy. This e–e term greatly distorts the charge density of the HF wave function, and the $N_{1,1}$ term repairs this, retrieving an additional 45%–50% of the DMC correlation energy. In the case of the more diffuse H_2 molecule the $N_{2,0}$ and $N_{1,1}$ terms have a different relative importance, but $J = N_{2,0} + N_{1,1}$ retrieves 70%–75% of the DMC correlation energy in the four molecules.

Like $N_{1,1}$, $N_{1,2}$ acts as a correction to the single-electron orbitals. This term provides no significant benefit in H_2O, where we have used high-quality orbitals, but it recovers 7% of the DMC correlation energy in H_2.

The effect of $N_{1,2}$ in N_2 is noteworthy in that the energy reduction obtained by adding this term to $J = N_{2,0} + N_{1,1}$ is about a factor of four times smaller than when added to the more accurate $J = N_{2,0} + N_{1,1} + N_{2,1}$. One would expect a term to retrieve more correlation energy when added to a smaller Jastrow factor, and this is the case for $N_{1,2}$ in the other molecules. We think that the distortion in the charge density caused by $N_{2,1}$ in N_2 is such that the single-electron correction effected by $N_{1,2}$ becomes more useful in its presence.

The $N_{2,1}$ term added to $J = N_{2,0} + N_{1,1} + N_{1,2}$ captures an additional 15%–20% of the DMC correlation energy. Higher-order terms added to $J = N_{2,0} + N_{1,1} + N_{2,1} + N_{1,2}$ yield significant gains in relative terms, with e–e–n–n terms retrieving 13% and 43% of the remaining DMC correlation energy remaining for N_2 and H_2, respectively, and the e–e–e–n term recovering 17% of the remaining DMC correlation energy.

E. Summary of results

Table III gives a comparison of the best single-determinant nonbackflow VMC energies we have found in the literature with those obtained in this work.

V. CONCLUSIONS

We have described a generalized Jastrow factor allowing terms that explicitly correlate the motions of n electrons with m static nuclei. These terms can be parameterized using various basis sets, including terms that involve dot products of interparticle position vectors. We have also introduced anisotropic cutoff functions. The formalism may be applied to systems with particle types and external potentials other than electrons and Coulomb potentials.

Optimization of the wave function is one of the most human- and computer-time consuming tasks in performing QMC calculations. We have performed term-by-term optimizations to understand how different terms in the Jastrow factor contribute to the electronic description of a system, and we hope that our analysis will serve as a guideline for constructing Jastrow factors for other systems.

We have tested these terms on HEGs, atoms, and molecules. The variational freedom from the higher-order terms generally improves the quality of the wave function. We have only considered single-determinant wave functions in this study, although our Jastrow factor can, of course, be used with other wave function forms.

We have demonstrated the construction and application of an e–e–n–n Jastrow factor term designed to describe van der Waals interactions between atoms. This term retrieves a large fraction of the van der Waals correlation energy in tests on the triplet state of H_2 at the proton separation that minimizes the total energy of the system.

We have found evidence for the importance of three-electron Jastrow terms in the low-density 1D and 2D HEGs. Improving the Jastrow factor for single-determinant backflow wave functions also leads to small improvements in the DMC energy of the 2D HEG. This demonstrates the indirect effect that improving the Jastrow factor can have on improving the nodal surface, as reported in Ref. [13].

We have made efforts to obtain accurate single-determinant VMC energies for most of the systems studied, but for BeH and N_2 we deliberately used inferior one-electron basis sets to see whether we could compensate for this with one-electron Jastrow terms. We find that this goal can be achieved by including an $N_{1,2}$ Jastrow term or anisotropic e–n terms, along with the usual $N_{1,1}$ term.

ACKNOWLEDGMENTS

We thank John Trail for producing the HF wave function for H_2. The calculations were performed on the Cambridge High Performance Computing Service facility.

APPENDIX A: IMPLEMENTATION

In this section we describe our design choices in implementing our Jastrow factor in the CASINO code [16]. The implementation principles are modularity and extensibility, embracing the flexibility that this Jastrow factor has by design rather than impeding it by focusing too strictly on performance.

1. Basis functions

The most important design requirement for modularity is that basis sets be dealt with separately rather than included in the Jastrow factor code. Basis functions, along with their derivatives when required, are evaluated and stored for later use. For any given term with expansion orders p and q there are $pN(N-1)/2$ e–e and qNM e–n functions to evaluate and store. Single-electron updates involve recalculating $p(N-1)$ and qM of these functions. Note that the number of basis functions that must be evaluated is independent of the term ranks n and m. Furthermore, it is possible to allow different terms to share basis functions that do not contain optimizable parameters so that, e.g., the natural powers involved in computing $N_{2,0}$ and $N_{1,1}$ can be reused for $N_{2,1}$.

A number of properties of the basis functions are required to construct the Jastrow factor. Equation (12) requires knowledge of whether basis functions are symmetric or antisymmetric, and their value, first radial derivative, and angular dependency at the origin are required by Eqs. (20), (21), (22), (23), (24), and (25). The one-contribution, no-contribution, and equal-product constraints in these equations require a table indexing distinct products of two basis functions at any value of their arguments. We implement interfaces that make these properties available so that basis functions can be treated as abstract objects in the construction of the Jastrow factor, which makes implementing new basis sets straightforward.

Cutoff functions are dealt with as additional basis sets with an expansion order of one, and we store information identifying the cutoff functions that are strictly zero, which we use to speed up evaluation of the Jastrow factor.

2. Evaluation of the Jastrow factor

For the evaluation of an arbitrary-rank Jastrow factor term it is necessary to use efficient procedures that iterate from a given set of electronic and nuclear indices $\mathbf{i}$ and $\mathbf{I}$ to the next in a specific order; explicit loops over scalar integer indices are not an option in static code since the loop depth is variable, and the memory usage of precomputing all possible $\mathbf{i}$ and $\mathbf{I}$ scales badly with system size for high ranks n and m. These procedures should take into account which cutoff functions are zero so that particle sets that do not contribute to the Jastrow factor are skipped. Efficient handling of localized Jastrow factor terms is important because it allows the cost of evaluating a term of any rank to scale linearly with system size if the cutoff lengths are held fixed. We implement a scheme where we construct a list of the electrons that are “connected” to each electron and each nucleus via nonzero cutoff functions. For the electronic indices, the value of i_1 is iterated between 1 and N, then the value of i_2 is iterated over the values in the list associated with electron i_1 that are greater than i_1, then the value of i_3 is iterated over the values in the intersection between the lists associated with i_1 and i_2 that are greater than i_2, and so on. The procedure that iterates over nuclear indices selects sets of nuclei whose connected-electron lists have nonzero intersections. We iterate over $\mathbf{I}$ in the outermost loop so that we can feed the intersection of all e–n lists into the $\mathbf{i}$ iterator as an initial list for index i_1.

The signature $\{\underline{\tilde{P}}(\mathbf{i}),\underline{\tilde{S}}(\mathbf{i},\mathbf{I})\}$ of each group of particles is computed inside the electronic and nuclear loops, identifying the linear parameter channel associated with the group of particles. We then loop over linear parameters in the channel, computing the products of the relevant basis functions which have already been precomputed. In terms without indexing constraints, consecutive linear parameters tend to have very similar expansion indices; i.e., they multiply most of the same basis functions. In order to save multiplications, it is convenient to buffer partial products so that, e.g., if only the last two of six expansion indices change from one parameter to the next, we can recover the product of the first four functions from the previous index set and save three of the five multiplications required to combine the six basis functions.

In typical QMC calculations individual electron moves, rather than full configuration moves, are proposed, which requires computing an acceptance probability involving the ratio of the trial wave function at the proposed and original positions. To calculate this efficiently one needs to compute the part of the Jastrow factor which depends on the position of a single particle i, ignoring the contributions not involving i. In our implementation we evaluate this one-electron Jastrow factor using Eq. (5), where we fix $i_1 = i$ and iterate over the rest of $\mathbf{i}$. The main difference from the evaluation of the full Jastrow factor is that $\mathbf{i}$ is not sorted, and the permutation required to sort $\mathbf{i}$, which amounts to inserting i at the correct position in $\mathbf{i}$, needs to be taken into account in the presence of antisymmetric e–e basis functions, since sign changes will be required by Eq. (12). The evaluation of gradients and Laplacians of the Jastrow factor term can be easily accommodated in the one-electron Jastrow evaluation code.

For performance reasons it is advisable to implement versions of the evaluation procedure for fixed ranks, with fixed-depth loops that can be optimized by compilers. We implement three optimized versions: one for e–e terms, one for e–n terms, and one for e–e–n terms. Other terms are handled by three generic procedures: one for terms without e–e functions, one for terms without e–n functions, and one for terms with both e–e and e–n functions.

APPENDIX B: TERM CONSTRUCTION EXAMPLE

Let us consider the $N_{3,0}$ term used for the 2D HEG in Sec. IV A2. This is a system with $N = 42$ electrons, half of each spin, and the $N_{3,0}$ term corresponds to $n = 3$, $m = 0$, expansion order $p = 4$, and basis functions

$$\Phi_\nu^P(\mathbf{r}) = r^{\nu-1}\left(\frac{L_P - r}{L_P}\right). \tag{B1}$$

The spin-pair dependency matrix $\underline{P}$ is of size $N \times N$, but in practice we specify a reduced 2×2 version of this matrix where each row (column) corresponds to different spins. For this system this matrix is

$$\underline{P} = \begin{pmatrix} 1 & 2 \\ 2 & 1 \end{pmatrix}, \tag{B2}$$

where we assign parallel-spin electron pairs a spin-pair dependency index of 1, and antiparallel-spin electron pairs an index of 2. The $\underline{P}$ matrix can be regarded as an intrinsic property of the system, but in some cases additional symmetries can be imposed in order to reduce the number of parameters in the Jastrow factor term; for example, we would achieve this by setting all elements of $\underline{P}$ to 1 in the 2D HEG, choosing to ignore the distinction between parallel- and antiparallel-spin electron pairs.

The elements in $\underline{P}$ determine pairwise properties and objects. For example, there are as many sets of nonlinear parameters in a Jastrow factor term as different values in $\underline{P}$; in this case, there are two cutoff lengths, L_1 for parallel-spin electron pairs and L_2 for antiparallel-spin electron pairs.

Four distinct types of three-electron groups can be formed: three up-spin electrons (↑↑↑), two up-spins and one down-spin (↑↑↓), one up-spin and two down-spins (↑↓↓), and three down-spin electrons (↓↓↓). The spin-pair dependency matrices for these four groups are

$$\underline{P}(\uparrow\uparrow\uparrow) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix};\ \underline{P}(\downarrow\downarrow\downarrow) = \begin{pmatrix} 0 & 1 & 1 \\ 1 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}, \tag{B3}$$

$$\underline{P}(\uparrow\uparrow\downarrow) = \begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix};\ \underline{P}(\uparrow\downarrow\downarrow) = \begin{pmatrix} 0 & 2 & 2 \\ 2 & 0 & 1 \\ 2 & 1 & 0 \end{pmatrix}. \tag{B4}$$

The matrices $\underline{P}(\uparrow\uparrow\uparrow)$ and $\underline{P}(\downarrow\downarrow\downarrow)$ are identical, and therefore correspond to a linear parameter channel which is used for groups of three electrons with parallel spins. The matrix $\underline{P}(\uparrow\downarrow\downarrow)$ can be transformed into $\underline{P}(\uparrow\uparrow\downarrow)$ via, e.g., the permutation matrix

$$\underline{U} = \begin{pmatrix} 0 & 0 & 1 \\ 0 & 1 & 0 \\ 1 & 0 & 0 \end{pmatrix}, \tag{B5}$$

and therefore both matrices correspond to a second linear parameter channel which is used for groups of three electrons with mixed spins. The signature of the first channel, which we refer to as the ↑↑↑ channel, is $\underline{P}(\uparrow\uparrow\uparrow)$, and the signature of the ↑↑↓ channel is $\underline{P}(\uparrow\uparrow\downarrow)$. Both of these matrices are considered sorted by our matrix-sorting algorithm.

The symmetry constraints for the parameters in each of the two channels depend on the above matrices. $\underline{P}(\uparrow\uparrow\uparrow)$ is invariant with respect to the application of any permutation, and for the ↑↑↑ channel Eq. (12) equates any two parameters with the same indices in a different order. The signature of the ↑↑↓ channel is invariant only with respect to one nontrivial permutation,

$$\underline{U} = \begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 0 \\ 0 & 0 & 1 \end{pmatrix}, \tag{B6}$$

and therefore the symmetry constraints for this channel are

$$\lambda^{\begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix}}_{\begin{pmatrix} 0 & \nu_{ij} & \nu_{ik} \\ \nu_{ij} & 0 & \nu_{jk} \\ \nu_{ik} & \nu_{jk} & 0 \end{pmatrix}} = \lambda^{\begin{pmatrix} 0 & 1 & 2 \\ 1 & 0 & 2 \\ 2 & 2 & 0 \end{pmatrix}}_{\begin{pmatrix} 0 & \nu_{ij} & \nu_{jk} \\ \nu_{ij} & 0 & \nu_{ik} \\ \nu_{jk} & \nu_{ik} & 0 \end{pmatrix}}. \tag{B7}$$

We do not allow expansion indices to be zero in the $N_{3,0}$ term, and the presence of cutoff functions prevents the "one contribution" constraint of Eq. (20) from being satisfied. It is therefore not possible to impose the Kato cusp conditions on the $N_{3,0}$ term. This is not a problem in practice because we use this term in conjunction with $N_{2,0}$.

Since the basis functions are isotropic, the only constraint applicable to this term at coalescence points is Eq. (22) for $l = m = 0$. The derivative of the radial projection of the basis function is

$$\left[\frac{\partial \mathcal{P}_{0,0}\left[\Phi_\nu^P(\mathbf{r})\right]}{\partial r}\right]_{r=0} = \delta_{\nu 2} - \delta_{\nu 1}\frac{C}{L_P}, \tag{B8}$$

and the constraint equation for index ν_{ij} is

$$\sum_{\nu_{ij}\nu_{ik}\nu_{jk}}^{\text{e.p.}} \left[\lambda^{\begin{pmatrix} 0 & P_{ij} & P_{ik} \\ P_{ij} & 0 & P_{jk} \\ P_{ik} & P_{jk} & 0 \end{pmatrix}}_{\begin{pmatrix} 0 & \nu_{ij} & \nu_{ik} \\ \nu_{ij} & 0 & \nu_{jk} \\ \nu_{ik} & \nu_{jk} & 0 \end{pmatrix}} \delta_{\nu_{ij}2} - \delta_{\nu_{ij}1}\frac{C}{L_{P_{ij}}}\right] = 0. \tag{B9}$$

Any two products of pairs of natural powers is equal if the sum of the exponents in each of them is equal. Therefore the "equal-product" constraint for the $N_{3,0}$ term is

$$\sum_{\nu_{ik}\nu_{jk}}^{\text{e.p.}} = \sum_{\nu_{ik}=1}^{\min(l-1,p)} \delta_{\nu_{jk},l-\nu_{ik}}, \tag{B10}$$

where l ranges from 2 to $2p$, and the constraint equation for index ν_{ij} thus reduces to

$$\sum_{\nu_{ik}=1}^{\min(l-1,p)} \left[\lambda \genfrac{}{}{0pt}{}{\begin{pmatrix} 0 & P_{ij} & P_{ik} \\ P_{ij} & 0 & P_{jk} \\ P_{ik} & P_{jk} & 0 \end{pmatrix}}{\begin{pmatrix} 0 & 2 & \nu_{ik} \\ 2 & 0 & l-\nu_{ik} \\ \nu_{ik} & l-\nu_{ik} & 0 \end{pmatrix}} - \lambda \genfrac{}{}{0pt}{}{\begin{pmatrix} 0 & P_{ij} & P_{ik} \\ P_{ij} & 0 & P_{jk} \\ P_{ik} & P_{jk} & 0 \end{pmatrix}}{\begin{pmatrix} 0 & 1 & \nu_{ik} \\ 1 & 0 & l-\nu_{ik} \\ \nu_{ik} & l-\nu_{ik} & 0 \end{pmatrix}} \frac{C}{L_{P_{ij}}} \right] = 0. \tag{B11}$$

The constraints for the three expansion indices are equal by symmetry in the ↑↑↑ channel, and there are two sets of constraints in the ↑↑↓ channel.

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

[2] J. Kolorenč and L. Mitas, Rep. Prog. Phys. **74**, 026502 (2011).

[3] B. M. Austin, D. Y. Zubarev, and W. A. Lester Jr., Chem. Rev. **112**, 263 (2011).

[4] J. B. Anderson, J. Chem. Phys. **65**, 4121 (1976).

[5] M. Casula, C. Attaccalite, and S. Sorella, J. Chem. Phys. **121**, 7110 (2004).

[6] P. López Ríos, A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs, Phys. Rev. E **74**, 066701 (2006).

[7] N. D. Drummond, M. D. Towler, and R. J. Needs, Phys. Rev. B **70**, 235119 (2004).

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

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

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

[11] M. Casula, S. Moroni, S. Sorella, and C. Filippi, J. Chem. Phys. **132**, 154113 (2010).

[12] C. Huang, C. J. Umrigar, and M. P. Nightingale, J. Chem. Phys. **107**, 3007 (1997).

[13] Y. Kwon, D. M. Ceperley, and R. M. Martin, Phys. Rev. B **48**, 12037 (1993).

[14] S. A. Alexander and R. L. Coldwell, J. Chem. Phys. **121**, 11557 (2004).

[15] M. Holzmann, B. Bernu, and D. M. Ceperley, Phys. Rev. B **74**, 104510 (2006).

[16] R. J. Needs, M. D. Towler, N. D. Drummond, and P. López Ríos, J. Phys.: Condens. Matter **22**, 023201 (2010).

[17] See Supplemental Material at http://link.aps.org/supplemental/10.1103/PhysRevE.86.036703 for the complete set of results.

[18] For example, $\sum_{i\neq j}^{N} f_{i,j}$ contains $N(N-1)$ terms, but if $f_{i,j} = f_{j,i}$, this operation can be simplified to $\sum_{i<j}^{N} f_{i,j}$, which contains $N(N-1)/2$ terms. In general, the sum of an object which is symmetric with respect to the interchange of any of its n indices, where these indices take nonrepeated values, can be rewritten as a sum over sorted sets of indices involving $n!$ times fewer terms than the original sum.

[19] Given a symmetric $n \times n$ matrix $\underline{S}$ and a rectangular $n \times m$ matrix $\underline{P}$, this algorithm finds the permutation of these matrices, $\tilde{\underline{P}} = \underline{U}\,\underline{P}\,\underline{U}^{\mathrm{T}}$ and $\tilde{\underline{S}} = \underline{U}\,\underline{S}\,\underline{R}^{\mathrm{T}}$, that correspond to the lowest collational order according to a predefined sorting criterion. The specifics of such criterion are unimportant; the only requirement is that for any $n \times n$ permutation matrix $\underline{A}$ and $m \times m$ permutation matrix $\underline{B}$, the application of the algorithm to the pair of matrices $\underline{A}\,\underline{P}\,\underline{A}^{\mathrm{T}}$ and $\underline{A}\,\underline{S}\,\underline{B}^{\mathrm{T}}$ also returns $\tilde{\underline{P}}$ and $\tilde{\underline{S}}$.

[20] M. C. Per, S. P. Russo, and I. K. Snook, J. Chem. Phys. **130**, 134103 (2009).

[21] M. C. Per, private communication.

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

[23] K. E. Schmidt and J. W. Moskowitz, J. Chem. Phys. **93**, 4172 (1990).

[24] C. Filippi and C. J. Umrigar, J. Chem. Phys. **105**, 213 (1996).

[25] P. Seth, P. López Ríos, and R. J. Needs, J. Chem. Phys. **134**, 084105 (2011).

[26] J. Toulouse and C. J. Umrigar, J. Chem. Phys. **126**, 084102 (2007).

[27] C. J. Umrigar, J. Toulouse, C. Filippi, S. Sorella, and R. G. Hennig, Phys. Rev. Lett. **98**, 110201 (2007).

[28] In general, attaining the DMC limit in VMC would require nonanalyticities in the Jastrow factor. Nonetheless this theoretical limit is useful for assessing the performance of a Jastrow factor.

[29] N. D. Drummond and R. J. Needs, Phys. Rev. Lett. **99**, 166401 (2007).

[30] R. M. Lee and N. D. Drummond, Phys. Rev. B **83**, 245114 (2011).

[31] N. D. Drummond and R. J. Needs, Phys. Rev. Lett. **102**, 126402 (2009).

[32] M. D. Brown, J. R. Trail, P. López Ríos, and R. J. Needs, J. Chem. Phys. **126**, 224110 (2007).

[33] C. Froese Fischer, G. Tachiev, G. Gaigalas, and M. R. Godefroid, Comput. Phys. Commun. **176**, 559 (2007).

[34] S. J. Chakravorty, S. R. Gwaltney, E. R. Davidson, F. A. Parpia, and C. F. Fischer, Phys. Rev. A **47**, 3649 (1993).

[35] J. Toulouse and C. J. Umrigar, J. Chem. Phys. **128**, 174101 (2008).

[36] D. Feller, K. A. Peterson, and D. A. Dixon, J. Chem. Phys. **129**, 204105 (2008).

[37] G. te Velde, F. M. Bickelhaupt, E. J. Baerends, C. Fonseca Guerra, S. J. A. van Gisbergen, J. G. Snijders, and T. Ziegler, J. Comput. Chem. **22**, 931 (2001).

[38] D. P. O'Neill and P. M. W. Gill, Mol. Phys. **103**, 763 (2005).

[39] R. Dovesi, V. R. Saunders, C. Roetti, R. Orlando, C. M. Zicovich-Wilson, F. Pascale, B. Civalleri, K. Doll, N. M. Harrison, I. J. Bush, P. D'Arco, and M. Llunell, CRYSTAL06 *User's Manual* (University of Torino, Torino, 2006).

[40] A. Ma, M. D. Towler, N. D. Drummond, and R. J. Needs, J. Chem. Phys. **122**, 224322 (2005).

[41] I. G. Gurtubay and R. J. Needs, J. Chem. Phys. **127**, 124306 (2007).

[42] B. K. Clark, M. A. Morales, J. McMinis, J. Kim, and G. E. Scuseria, J. Chem. Phys. **135**, 244105 (2011).

[43] W. Kolos and L. Wolniewicz, J. Chem. Phys. **43**, 2429 (1965).

[44] L. Laaksonen, P. Pyykkö, and D. Sundholm, Comput. Phys. Rep. **4**, 313 (1986).

[45] G. Staszewska and L. Wolniewicz, J. Mol. Spectrosc. **198**, 416 (1999).

[46] N. Nemec, M. D. Towler, and R. J. Needs, J. Chem. Phys. **132**, 034111 (2010).