Multi-Jastrow trial wavefunctions for electronic structure calculations with quantum Monte Carlo

Thomas Bouabça,[1] Benoît Braïda,[2] and Michel Caffarel[1,a)]

[1]*Laboratoire de Chimie et Physique Quantiques, CNRS UMR5626-IRSAMC et Université de Toulouse, Toulouse Cedex 31000, France*
[2]*Laboratoire de Chimie Théorique, CNRS UMR7616 et Université Pierre et Marie Curie, Case Courrier 137, 4 Place Jussieu, 75252 Paris, France*

A new type of electronic trial wavefunction suitable for quantum Monte Carlo calculations of molecular systems is presented. In contrast with the standard Jastrow–Slater form built with a unique global Jastrow term, it is proposed to introduce individual Jastrow factors attached to molecular orbitals. Such a form is expected to be more physical since it allows to describe *differently* the local electronic correlations associated with various molecular environments ($1s$-core orbitals, $3d$-magnetic orbitals, localized two-center σ-orbitals, delocalized π-orbitals, atomic lone pairs, etc.). In contrast with the standard form, introducing different Jastrow terms allows us to change the nodal structure of the wavefunction, a point which is important in the context of building better nodes for more accurate fixed-node diffusion Monte Carlo (FN-DMC) calculations. Another important aspect resulting from the use of local Jastrow terms is the possibility of defining and preoptimizing local and *transferable* correlated units for building complex trial wavefunctions from simple parts. The practical aspects associated with the computation of the intricate derivatives of the multi-Jastrow trial function are presented in detail. Some first illustrative applications for atoms of increasing size (O, S, and Cu) and for the potential energy curve and spectroscopic constants of the FH molecule are presented. In the case of the copper atom, the use of the multi-Jastrow form at the variational Monte Carlo level has allowed us to improve significantly the value of the total ground-state energy (about 75% of the correlation energy with only one determinant and three atomic orbital Jastrow factors). In the case of the FH molecule (fluorine hydride), it has been found that the multi-Jastrow nodes lead to an almost exact FN-DMC value of the dissociation energy [$D_0 = -140.7(4)$ kcal/mol instead of the estimated nonrelativistic Born–Oppenheimer exact value of -141.1], which is not the case with standard nodes, $D_0 = -138.3(4)$ kcal/mol. © *2010 American Institute of Physics*. [doi:10.1063/1.3457364]

I. INTRODUCTION

At the heart of quantum Monte Carlo (QMC) simulations of electronic ground-state properties is the use of a trial wavefunction, Ψ_T, approximating the unknown wavefunction, Ψ_0. The role played by Ψ_T is central and many-fold. First, it is used to implement the "importance sampling" strategy common to any *efficient* Monte Carlo scheme (either classical or quantum). Indeed, apart for very simple systems, Monte Carlo estimators are expected to converge only if the random sampling of the underlying configuration space is made efficiently, that is to say, only regions associated with a significant probability density are visited. Second, the trial wavefunction is essential for reducing the statistical fluctuations on the total energy. This property is based on the existence of a so-called zero-variance (ZV) principle which states that the closer the trial wavefunction is from the exact wavefunction, the smallest the statistical errors on the energy are. In the limit, $\Psi_T = \Psi_0$ the statistical error entirely vanishes (ZV). Note that this ZV principle valid for the energy has

been recently extended to arbitrary properties.[1] Finally, a last but fundamental role played by the trial wavefunction in the case of fermionic systems is its role in imposing the Pauli exclusion principle. As well-known, there does not exist a stable and exact QMC algorithm for fermionic systems (because of the famous "sign problem," for molecules see, e.g., Ref. 2). In practice, this difficulty is circumvented by using a stable algorithm consisting in solving the Schrödinger equation with the additional constraint that the solution has the same sign as the trial wavefunction. Unfortunately, this constraint which implies to impose to Ψ_0 the same nodal hypersurfaces as Ψ_T introduces a small bias, the so-called fixed-node (FN) error. This error is, in general, a very small fraction of the total energy; nevertheless, it can lead to undesirable physical consequences when considering the small energy differences we are interested in (electron affinities, ionization potentials, binding energies, etc.). Note also that it can be shown that the FN energy is an upper bound of the exact energy (variational principle[3]).

Now, in view of the practical importance of all these points, an intense activity has been developed toward

a)Author to whom correspondence should be addressed. Electronic mail: caffarel@irsamc.ups-tlse.fr.

devising better trial wavefunctions associated with smaller statistical fluctuations (to decrease the computational effort and/or to treat larger systems) and smaller FN errors (better control of the quality of the results). The general guide for improving trial wavefunctions is to use functional forms which incorporate as much as possible the various physical and mathematical properties of the exact unknown wavefunction. In addition, the form of the trial wavefunctions must absolutely be in practice as compact as possible since the function, its first, and second derivatives are to be calculated at each Monte Carlo step. Finally, on top of that, it is desirable to optimize by minimization of the energy or the variance all or part of the parameters appearing in the trial wavefunction.

The most standard form used presently for the trial wavefunction, known as the Slater–Jastrow form, is written as

$$\Psi_T = e^{J(\mathbf{r}_1,\ldots,\mathbf{r}_N)} \sum_{k=1}^{N_{\text{det}}} c_k \, \text{Det}_k(\{\Phi_i^\uparrow\}) \text{Det}_k(\{\Phi_i^\downarrow\}), \tag{1}$$

with N being the number of electrons, N_{det} is the number of terms in the multideterminant expansion, and $\{\Phi_i^\sigma\}(\sigma=\uparrow,\downarrow)$ is a set of molecular orbitals. In this expression the prefactor e^J is known as the Jastrow factor. Its role is to impose the exact behavior of the wavefunction in the $[r_{ij} \to 0]$-limit (electron-electron cusp condition) and, also, to incorporate some two-body (electron-electron and electron-nucleus) and three-body (electron-electron-nucleus) correlations (to describe the best as possible the shape of the Coulomb hole[4]). The determinantal part is, in general, issued from some *ab initio* wavefunction or density functional theory (DFT) calculations. The multideterminant expansion is introduced for describing situations where the exact wavefunction is expected to be strongly multiconfigurational (bond-breaking processes, excited-states, etc.). Because molecular orbitals are optimized within a pure *ab initio* framework without the presence of a Jastrow term, it is, in general, desirable to reoptimize most of the parameters of the trial wavefunction[5] (Jastrow parameters, molecular orbitals, and expansion coefficients c_k).

Other forms for the trial wavefunction have been introduced in literature but, so far, they have remained of marginal use. We can cite, for example, the geminal wavefunction of Sorella and co-workers,[6] the Pfaffian wavefunction of Mitáš and collaborators,[7] and the backflow (BF) trial wavefunction of López Ríos *et al.*[8]

Here, we propose a new trial wavefunction valid for arbitrary molecular systems. This wavefunction is a generalization of the standard form, Eq. (1), where the common Jastrow factor is replaced by several Jastrow factors e^{J_i}, each of them being attached to a particular molecular orbital, Φ_i. In what follows, this trial wavefunction will be called the multi-Jastrow wavefunction. The expression of the new wavefunction is written as follows:

$$\Psi_T = \sum_{k=1}^{N_{\text{det}}} c_k \, \text{Det}_k(\{e^{J_i}\Phi_i^\uparrow\}) \text{Det}_k(\{e^{J_i}\Phi_i^\downarrow\}). \tag{2}$$

Note that in this latter formula the general case corresponding to the use of a different Jastrow term for each orbital is considered. However, as we shall see in the applications, using as many Jastrow terms as orbitals is, in general, not necessary: We can avoid introducing Jastrow terms for orbitals which do not play a critical role and/or use a common Jastrow term for molecular orbitals equivalent by symmetry. The physical idea behind this form is to allow an *orbital-dependent* description of the structure of the correlation effects. Indeed, it is quite legitimate not to place on the same footing the correlation effects between electrons occupying a $1s$-core orbital, a $3d$-magnetic orbital, an atomic lone pair, a localized two-center σ-orbital, a delocalized π-orbital, etc. As we shall see in our applications, the multi-Jastrow trial wavefunction appears to be a physical improvement with respect to the standard form, Eq. (1). Not only the variational energy is lowered but also the change of the nodal structure associated with the introduction of local Jastrow terms into determinants can also lead to better results. Finally, employing a multi-Jastrow form opens perspectives toward devising systematic procedures for constructing good-quality wavefunctions for electronic systems of arbitrary size.

The organization of this paper is as follows. In Sec. II we present the multi-Jastrow trial wavefunction. In Sec. III the algorithms for optimizinig the trial wavefunction are briefly presented. In Sec. IV the most difficult technical part of the approach, namely, evaluating efficiently the derivatives of the multi-Jastrow trial wavefunction with respect to spatial coordinates and parameters, is presented in detail. In Sec. V, illustrative applications for the oxygen, sulfur, and copper atoms and the FH molecule are presented. In Sec. VI the computational cost associated with the use of multi-Jastrow wavefunctions is discussed. Finally, we present some summary and perspectives in the Sec. VII.

II. THE MULTI-JASTROW TRIAL WAVEFUNCTION

A. The standard form

As briefly recalled in Sec. I the standard form used for the trial wavefunction is as follows:

$$\Psi_T = e^{J(\mathbf{r}_1,\ldots,\mathbf{r}_N)} \sum_{k=1}^{N_{\text{det}}} c_k \begin{vmatrix} \phi_{k_1^\uparrow}(\mathbf{r}_1) & \ldots & \phi_{k_1^\uparrow}(\mathbf{r}_{N_\uparrow}) \\ \vdots & \vdots & \vdots \\ \phi_{k_{N_\uparrow}^\uparrow}(\mathbf{r}_1) & \ldots & \phi_{k_{N_\uparrow}^\uparrow}(\mathbf{r}_{N_\uparrow}) \end{vmatrix}$$
$$\times \begin{vmatrix} \phi_{k_1^\downarrow}(\mathbf{r}_{N_\uparrow+1}) & \ldots & \phi_{k_1^\downarrow}(\mathbf{r}_N) \\ \vdots & \vdots & \vdots \\ \phi_{k_{N_\downarrow}^\downarrow}(\mathbf{r}_{N_\uparrow+1}) & \ldots & \phi_{k_{N_\downarrow}^\downarrow}(\mathbf{r}_N) \end{vmatrix}. \tag{3}$$

In this expression the prefactor e^J is the Jastrow factor. Note that for the sake of simplicity J or e^J will be called here Jastrow factors indifferently. Many different forms for J have been proposed in literature. A typical "minimal" form is given as follows:

$$J(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \sum_{i<j} \frac{a_{\sigma_{ij}} r_{ij}}{1 + b_{\sigma_{ij}} r_{ij}} - \sum_{\alpha} c_{\alpha} \sum_i r_{i\alpha}, \tag{4}$$

where the Latin indices i and j refer to electrons and the Greek index α to nuclei. Parameters $a_{\sigma_{ij}}$ are introduced to impose the electron-electron cusp condition, namely, $a_{\sigma_{ij}} = 1/2$ for spin-unlike electrons and $a_{\sigma_{ij}} = 1/4$ for spin-like electrons. The parameters $b_{\sigma_{ij}}$ are introduced to take into account the screening of the electron-electron interaction at large interelectronic distances, and the parameters c_{α} are introduced to allow some readjustment of the electronic density once the r_{ij} terms have been introduced into the wavefunction. More sophisticated forms can be written and are generally expressed under the generic form,

$$J(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \sum_{i<j} v_{e-e}(r_{ij}) + \sum_{\alpha} v_{e-n}(r_{i\alpha})$$
$$+ \sum_{i<j} \sum_{\alpha} v_{e-e-n}(r_{ij}, r_{i\alpha}, r_{j\alpha}), \tag{5}$$

where the functions, v_{e-e}, v_{e-n}, and v_{e-e-n} describe the two-body electron-electron, electron-nucleus correlations, and the three-body electron-electron-nucleus correlations, respectively. Each of these functions is usually expanded in terms of simple elementary functions (polynomials, Padé, sums of exponentials, etc.).

Here, we shall use the following form:

$$J(\mathbf{r}_1, \ldots, \mathbf{r}_N) = \frac{1}{2} \sum_{j \neq i} \sum_{\alpha} [s_{\sigma_{ij}}(\tilde{r}_{ij}) - p_{\alpha}(\tilde{r}_{i\alpha}) - p_{\alpha}(\tilde{r}_{j\alpha})$$
$$+ g^{(1)}_{\alpha\sigma_{ij}} \tilde{r}_{i\alpha}^2 \tilde{r}_{j\alpha}^2 + g^{(2)}_{\alpha\sigma_{ij}} (\tilde{r}_{i\alpha}^2 + \tilde{r}_{j\alpha}^2) \tilde{r}_{ij}^2], \tag{6}$$

where $\tilde{r}_{i\alpha}$ and $\tilde{r}_{ij}$ are renormalized electron-nucleus and interelectronic distances, respectively, whose expressions are given by

$$\tilde{r}_{ij} \equiv \frac{r_{ij}}{1 + b_{\sigma_{ij}} r_{ij}},$$

$$\tilde{r}_{i\alpha} \equiv \frac{r_{i\alpha}}{1 + b_{\alpha} r_{i\alpha}},$$

and where the functions $s_{\sigma_{ij}}(r)$ and $p_{\alpha}(r)$ are expressed as

$$s_{\sigma_{ij}}(r) = \frac{1}{N_{\text{nucl}}} (e^{(1)}_{\sigma_{ij}} r + e^{(2)}_{\sigma_{ij}} r^2 + e^{(3)}_{\sigma_{ij}} r^3 + e^{(4)}_{\sigma_{ij}} r^4),$$

$$p_{\alpha}(r) = \frac{1}{N_{\text{nucl}}} (f^{(1)}_{\alpha} r + f^{(2)}_{\alpha} r^2 + f^{(3)}_{\alpha} r^3 + f^{(4)}_{\alpha} r^4).$$

In these formulas N_{nucl} denotes the number of nuclei and the quantities $\{b_{\sigma_{ij}}, b_{\alpha}, e^{(l)}_{\sigma_{ij}}, f^{(l)}_{\alpha}, g^{(l)}_{\alpha\sigma_{ij}}\}$ play the role of parameters (there are $10 + 9N_{\text{nucl}}$ such parameters).

In the expression, Eq. (3), of the trial wavefunction the determinantal part is a finite sum of N_{det} Slater determinants of single-particle molecular orbitals, c_k being the various amplitude coefficients. $N_{\uparrow}$ and $N_{\downarrow}$ are the number of $\uparrow$ (or α) and $\downarrow$ (or β) electrons, respectively. This part is obtained using standard DFT or *ab initio* wavefunction-based method: Hartree–Fock (HF), complete active space self-consistent field (CASSCF), small-configuration interaction, etc. The various molecular orbitals $\phi_{k_i^{\sigma}}$ ($\sigma = \uparrow, \downarrow$) are extracted from a set of active orbitals $\phi_1(\mathbf{r}), \ldots, \phi_{N_{\text{act}}}(\mathbf{r})$. Note that the formalism used in QMC calculations is a spin-free formalism.[9] The trial wavefunction depends only on space coordinates, $\mathbf{r}_i$, and, the molecular orbitals employed are spatial orbitals depending only on the ordinary three-dimensional (3D)-coordinates. To get a wavefunction obeying the Pauli principle, the antisymmetry of the wavefunction under the exchange of the spatial coordinates of any pair of electrons having the same spin has to be imposed.[9] Here, this property is fulfilled because the global Jastrow term is by construction fully symmetric under the exchange of any pair of electrons and because the product of two Slater determinants for $\sigma = \uparrow$ and $\sigma = \downarrow$ realizes the antisymmetry.

B. The multi-Jastrow form

The multi-Jastrow form is similar to the standard form, except that the global Jastrow term is no longer attached to the determinantal part as a whole like in Eq. (3), but is now split into multiple Jastrow terms e^{J_i} attached to each molecular orbital $\phi_i(\mathbf{r})$. In what follows, these individual Jastrow terms will be referred to as *local* Jastrow terms. The denomination local is chosen here to emphasize on the fact that the Jastrow term will have a non-negligible impact only in the region of space where the (atomic or molecular) orbital will take its main values. To be more precise, the wavefunction proposed here is given as

$$\Psi_T = \sum_{k=1}^{N_{\text{det}}} c_k \begin{vmatrix} e^{J_{k_1^{\uparrow}}(\mathbf{r}_1|\mathbf{r}_i \neq \mathbf{r}_1)} \phi_{k_1^{\uparrow}}(\mathbf{r}_1) & \cdots & e^{J_{k_1^{\uparrow}}(\mathbf{r}_{N_{\uparrow}}|\mathbf{r}_i \neq \mathbf{r}_{N_{\uparrow}})} \phi_{k_1^{\uparrow}}(\mathbf{r}_{N_{\uparrow}}) \\ \vdots & \vdots & \vdots \\ e^{J_{k_{N_{\uparrow}}^{\uparrow}}(\mathbf{r}_1|\mathbf{r}_i \neq \mathbf{r}_1)} \phi_{k_{N_{\uparrow}}^{\uparrow}}(\mathbf{r}_1) & \cdots & e^{J_{k_{N_{\uparrow}}^{\uparrow}}(\mathbf{r}_{N_{\uparrow}}|\mathbf{r}_i \neq \mathbf{r}_{N_{\uparrow}})} \phi_{k_{N_{\uparrow}}^{\uparrow}}(\mathbf{r}_{N_{\uparrow}}) \end{vmatrix} \begin{vmatrix} e^{J_{k_1^{\downarrow}}(\mathbf{r}_{N_{\uparrow}+1}|\mathbf{r}_i \neq \mathbf{r}_{N_{\uparrow}+1})} \phi_{k_1^{\downarrow}}(\mathbf{r}_{N_{\uparrow}+1}) & \cdots & e^{J_{k_1^{\downarrow}}(\mathbf{r}_N|\mathbf{r}_i \neq \mathbf{r}_N)} \phi_{k_1^{\downarrow}}(\mathbf{r}_N) \\ \vdots & \vdots & \vdots \\ e^{J_{k_{N_{\downarrow}}^{\downarrow}}(\mathbf{r}_{N_{\uparrow}+1}|\mathbf{r}_i \neq \mathbf{r}_{N_{\uparrow}+1})} \phi_{k_{N_{\downarrow}}^{\downarrow}}(\mathbf{r}_{N_{\uparrow}+1}) & \cdots & e^{J_{k_{N_{\downarrow}}^{\downarrow}}(\mathbf{r}_N|\mathbf{r}_i \neq \mathbf{r}_N)} \phi_{k_{N_{\downarrow}}^{\downarrow}}(\mathbf{r}_N) \end{vmatrix}. \tag{7}$$

In this expression the set of orbitals $\{\phi_{k_i^\sigma}\}$, $\sigma=\uparrow,\downarrow$ of a given component of the sum is a subset of size N_σ of the full set of the active molecular orbitals used. The local Jastrow functions $J_k(\mathbf{r}_j|\mathbf{r}_i \neq \mathbf{r}_j)$ are arbitrary (smooth) functions of *all* electronic coordinates $(\mathbf{r}_1, \ldots, \mathbf{r}_N)$. For convenience, we have adopted the following notation:

$$J_k(\mathbf{r}_j|\mathbf{r}_i \neq \mathbf{r}_j) \equiv J_k(\mathbf{r}_j|\mathbf{r}_1, \ldots, \mathbf{r}_i \neq \mathbf{r}_j, \ldots, \mathbf{r}_N). \tag{8}$$

Remark that in the list of arguments of the local Jastrow terms the "reference" electron (here, labeled with j) is the electron occupying the molecular orbital attached to the local Jastrow and has been particularized from the $(N-1)$ remaining "external" electrons. This distinction is introduced because we impose to all local Jastrows J_k to be symmetric under the exchange of any pair of external electrons. This property can be expressed as follows:

$$J_k(\mathbf{r}_j|\mathbf{r}_1, \ldots, \mathbf{r}_l, \ldots, \mathbf{r}_m, \ldots, \mathbf{r}_N)$$
$$= J_k(\mathbf{r}_j|\mathbf{r}_1, \ldots, \mathbf{r}_m, \ldots, \mathbf{r}_l, \ldots, \mathbf{r}_N) \tag{9}$$
$$\forall l \neq j \quad \text{and} \quad \forall m \neq j.$$

Now, the elementary but most important point to remark is that our proposed form (7) still obeys the Pauli exclusion principle. To verify this we just need to check that the wavefunction changes its sign when two electrons having the same spin are exchanged. Without loss of generality, let us consider that the two electrons to be exchanged are electrons 1 and 2 of spin $\uparrow$. When doing the exchange, the first and second columns of all $\uparrow$-determinants are exchanged and the signs of these determinants are changed. In all other columns (columns of $\uparrow$-determinants except columns 1 and 2 and all columns of $\downarrow$-determinants), electrons 1 and 2 have the status of external electrons. Now, because of the symmetry property, Eq. (9), this exchange leaves the wavefunction unchanged. Finally, as desired, the overall sign of the wavefunction is changed when spinlike electrons are exchanged.

Various forms for the local Jastrow terms can be used. In this work, in order to allow the fairest possible comparison between results obtained with our new proposal and with standard wavefunctions using a single common Jastrow, the same functional form for both types of Jastrow has been chosen. Accordingly, for each Jastrow J_k we define the renormalized interelectronic, $\tilde{r}_{ij}^{(k)}$, and electron-nucleus distances $\tilde{r}_{i\alpha}^{(k)}$ defined as

$$\tilde{r}_{ij}^{(k)} \equiv \frac{r_{ij}}{1+b_k r_{ij}},$$

$$\tilde{r}_{i\alpha}^{(k)} \equiv \frac{r_{i\alpha}}{1+b_{\alpha k} r_{i\alpha}},$$

and the local Jastrow is written as

$$J_k(\mathbf{r}_i|\mathbf{r}_j \neq \mathbf{r}_i) = \frac{1}{2}\sum_{j \neq i}\sum_\alpha [s_k(\tilde{r}_{ij}) - p_{k\alpha}(\tilde{r}_{i\alpha}) - p_{k\alpha}(\tilde{r}_{j\alpha})$$
$$+ g_{k\alpha}^{(1)}\tilde{r}_{i\alpha}^2\tilde{r}_{j\alpha}^2 + g_{k\alpha}^{(2)}(\tilde{r}_{i\alpha}^2 + \tilde{r}_{j\alpha}^2)\tilde{r}_{ij}^2], \tag{10}$$

where

$$s_k(r) = \frac{1}{N_{\text{nucl}}}(e_k^{(1)}r + e_k^{(2)}r^2 + e_k^{(3)}r^3 + e_k^{(4)}r^4),$$

$$p_{k\alpha}(r) = \frac{1}{N_{\text{nucl}}}(f_{k\alpha}^{(1)}r + f_{k\alpha}^{(2)}r^2 + f_{k\alpha}^{(3)}r^3 + f_{k\alpha}^{(4)}r^4).$$

In contrast with the standard Jastrow factor, note that the parameters b_k, $e_k^{(l)}$, and $g_{k\alpha}^{(l)}$ do not depend on the spinlike or spin-unlike character of the electron pair. This latter property is primordial to have a trial wavefunction fulfilling the antisymmetry property of fermions. In these expressions, the quantities $\{b_k, b_{\alpha k}, e_k^{(l)}, f_{k\alpha}^{(l)}, g_{k\alpha}^{(l)}\}$ play the role of parameters. For a given local Jastrow, there are $5+7N_{\text{nucl}}$ such parameters.

Now, a few important remarks are in order. First, as already briefly discussed in Sec. I, the physical role played by the local Jastrow terms is rather clear. Each reference electron occupying a given molecular orbital interacts with the complete set of the $(N-1)$ remaining external electrons. The nature of this interaction and its representation via a Jastrow term is clearly very dependent on the type of molecular orbital occupied by the reference electron. For example, $1s$ electrons close to some nucleus (large-Z attraction) do not experience the same local Coulombic interaction as $3d$ electrons moving in a small region of space crowded by many electrons (large electron-electron repulsion). This is also the case when considering lone pair electrons which are of atomic nature or π-electrons involved in a multiple bond which can be mainly involved between two atoms or circulating over a ring (aromatic compounds) and so on. A second remark is that to build a properly antisymmetrized wavefunction constructed from fully correlated elementary blocks (here, the product $e^{J_i}\phi_i$ depending explicitly on *all* electronic coordinates) is, in general, not tractable from a computational point of view because of the $N!$ problem coming from the $\uparrow$- and $\downarrow$-antisymmetrizers (sums over all possible permutations). Here, the particular form of our trial wavefunction where a full symmetry between external electrons coordinates has been imposed allows us to re-express the antisymmetrizers as standard Slater determinants like in the usual case where the antisymmetrizers are applied on a product of one-particle functions. Related to this remark, let us note that it would have been more physical to consider a spin-dependent interaction between the reference electron of a given spin and external electrons of the same and opposite spin. Unfortunately, introducing such a structure would have led to untractable form for the trial wavefunction.

Finally, a last remark concerns the fact that introducing orbitals which depend on the positions of more than one-electron as done here is not new. In the case of the geminal wavefunction of Sorella and co-workers,[6] or the Pfaffian wavefunction of Mitáš,[7] the one-electron orbitals are replaced by a two-electron function with the purpose of better reproducing the electron pair correlations. In the case of the BF-type trial wavefunctions (see, e.g., Ref. [8]) molecular orbitals depending on the entire set of electron-coordinates via the use of a so-called BF displacement are introduced. How-

ever, despite this common feature the differences with our multi-Jastrow form are important and are briefly discussed in the next section.

C. Relationship with BF-type trial wavefunction

A multideterminantal BF trial wavefunction can be obtained from the standard Slater–Jastrow form, Eq. (3), by

adding to the coordinates $\mathbf{r}_i$ of the particles in the Slater a BF displacement depending on the coordinates of all particles,

$$\mathbf{r}_i \rightarrow \mathbf{r}_i + \boldsymbol{\xi}_i(\mathbf{r}_1, \ldots, \mathbf{r}_N). \tag{11}$$

Physically, this BF displacement is supposed to reproduce the perturbation of the flow of particles around a given particle. Using the same notations as above, the BF wavefunction can be written as follows:

$$
\Psi_T = e^{J(\mathbf{r}_1, \ldots, \mathbf{r}_N)} \sum_{k=1}^{N_{\text{det}}} c_k
\begin{vmatrix}
\phi_{k_1^\uparrow}(\mathbf{r}_1 + \boldsymbol{\xi}_1(\mathbf{r}_1, \ldots, \mathbf{r}_N)) & \ldots & \phi_{k_1^\uparrow}(\mathbf{r}_{N_\uparrow} + \boldsymbol{\xi}_{N_\uparrow}(\mathbf{r}_1, \ldots, \mathbf{r}_N)) \\
\vdots & \vdots & \vdots \\
\phi_{k_{N_\uparrow}^\uparrow}(\mathbf{r}_1 + \boldsymbol{\xi}_1(\mathbf{r}_1, \ldots, \mathbf{r}_N)) & \ldots & \phi_{k_{N_\uparrow}^\uparrow}(\mathbf{r}_{N_\uparrow} + \boldsymbol{\xi}_{N_\uparrow}(\mathbf{r}_1, \ldots, \mathbf{r}_N))
\end{vmatrix}
$$

$$
\times
\begin{vmatrix}
\phi_{k_1^\downarrow}(\mathbf{r}_{N_\uparrow+1} + \boldsymbol{\xi}_{N_\uparrow+1}(\mathbf{r}_1, \ldots, \mathbf{r}_N)) & \ldots & \phi_{k_1^\downarrow}(\mathbf{r}_N + \boldsymbol{\xi}_N(\mathbf{r}_1, \ldots, \mathbf{r}_N)) \\
\vdots & \vdots & \vdots \\
\phi_{k_{N_\downarrow}^\downarrow}(\mathbf{r}_{N_\uparrow+1} + \boldsymbol{\xi}_{N_\uparrow+1}(\mathbf{r}_1, \ldots, \mathbf{r}_N)) & \ldots & \phi_{k_{N_\downarrow}^\downarrow}(\mathbf{r}_N + \boldsymbol{\xi}_N(\mathbf{r}_1, \ldots, \mathbf{r}_N))
\end{vmatrix}. \tag{12}
$$

Comparing Eqs. (7) and (12) allows us to see the differences between how the electronic correlations are introduced. In the BF case, the local displacement ξ for a given electron is independent on the orbital occupied. In contrast, in the multi-Jastrow case the coupling with other electrons is strongly dependent on the orbital occupied, a point which is central in our approach. Another point is that in the multi-Jastrow case the partitioning of the 3D-space between different regions—atomic regions, bonding regions, etc.—is possible by using localized orbitals. Furthermore, it is not necessary to introduce a Jastrow term for each molecular orbital, a point which is not possible with the BF form. Hence, our multi-Jastrow form appears to be much more flexible than the BF wavefunction.

III. OPTIMIZATION OF THE TRIAL WAVEFUNCTION

In this section we present the optimization scheme employed in this work. As known, to define an efficient procedure for optimizing the parameters entering QMC wavefunctions is, in general, nontrivial. For example, we have found that the standard variance-minimization algorithm using a correlated approach of Umrigar et al.[10] was insufficient to get reasonable parameters for our multi-Jastrow wavefunctions. We have therefore turned our attention to the very recently proposed method of Umrigar and collaborators,[11] based on the use of a linear expansion of the trial wavefunction within the space spanned by the wavefunction and its first derivatives with respect to all parameters,

$$\psi_T^{\text{lin}}(\mathbf{p}, \mathbf{R}) = \psi_T(\mathbf{p}_0, \mathbf{R}) + \sum_{i=1}^{N_{\text{opt}}} (\mathbf{p} - \mathbf{p}_0)_i \frac{\partial \psi_T}{\partial p_i}(\mathbf{p}_0, \mathbf{R}), \tag{13}$$

where $\mathbf{R} \equiv (\mathbf{r}_1, \ldots, \mathbf{r}_N)$ and $\mathbf{p}$ is the set of the N_{opt} parameters

to be optimized. The optimal parameters (lowest total energy) are obtained by diagonalizing the Hamiltonian within the linear functional space consisting of $\psi_0(\mathbf{R}) \equiv \psi_T(\mathbf{p}_0, \mathbf{R})$ and the N_{opt} functions $\psi_i(\mathbf{R}) \equiv (\partial \psi_T / \partial p_i)(\mathbf{p}_0, \mathbf{R})$. The ($N_{\text{opt}}$ + 1)-dimensional generalized eigenvalue problem is thus written as

$$\mathbf{H} \Delta \mathbf{p} = E \mathbf{S} \Delta \mathbf{p}, \tag{14}$$

where $\Delta \mathbf{p}_i = (\mathbf{p} - \mathbf{p}_0)_i$ for $i = 1$ to N_{opt} and $\Delta \mathbf{p}_0 = 1$. The matrix elements are given by $H_{ij} = \langle \psi_i | H | \psi_j \rangle$ and $S_{ij} = \langle \psi_i | \psi_j \rangle$. In practice, this approach has been implemented by using the same strategy as in the original work,[11] namely,

(i)   calculation of the matrix elements H_{ij} and S_{ij} using the variational Monte Carlo (VMC) approach with $\psi_T^2(\mathbf{p}_0, \mathbf{R})$ as probability density, the estimators of the matrix elements obeying a "strong" ZV principle as introduced in Ref. 12;

(ii) renormalization of the components of $\Delta \mathbf{p}$ corresponding to nonlinear parameters by using a rescaling factor;

(iii) use of a stabilization procedure to minimize the effect of the statistical noise.

The preceding steps are iterated until the change of parameters, $\Delta \mathbf{p}$'s, is negligible. In this procedure only the last step has been modified here and it has been done as follows. As proposed in Ref. 11 a positive constant, λ, is added to the diagonal part of the Hamiltonian matrix except for the first element,

$$H_{ij}(\lambda) \equiv H_{ij} + \lambda \delta_{ij}(1 - \delta_{i0}). \tag{15}$$

Physically, λ can be viewed as a repulsive potential making the components $\Delta p_i(\lambda)$ as small as desired when the magni-

tude of λ is increased. Now, let us denote the ground-state eigensolutions of the eigenproblem associated with $H(\lambda)$ as $\psi_T^{\text{lin}}(\lambda)$ and $E(\lambda)$. Here, we propose to use the wavefunction $\psi_T^{\text{lin}}(\lambda)$ as a trial wavefunction for the true Hamiltonian H (corresponding to $\lambda=0$). Doing that, it is easy to verify that the corresponding variational energy, $E^{\text{lin}}(\lambda)$, is given by

$$E^{\text{lin}}(\lambda) = \frac{\langle \psi_T^{\text{lin}}(\lambda)|H|\psi_T^{\text{lin}}(\lambda)\rangle}{\langle \psi_T^{\text{lin}}(\lambda)|\psi_T^{\text{lin}}(\lambda)\rangle} = E(\lambda) - \lambda \frac{\sum_{i=1}^{N_{\text{opt}}}\Delta p_i^2(\lambda)}{1 + \sum_{i=1}^{N_{\text{opt}}}\Delta p_i^2(\lambda)}. \tag{16}$$

To get the optimal value of λ the variational energy, $E^{\text{lin}}(\lambda)$, is minimized. In practice, we evaluate the variational energy, $\langle \psi_T^{\text{lin}}(\lambda)|H|\psi_T^{\text{lin}}(\lambda)\rangle/\langle \psi_T^{\text{lin}}(\lambda)|\psi_T^{\text{lin}}(\lambda)\rangle$, for different values of λ with a correlated approach using the initial density as reference density,

$$E^{\text{lin}}(\lambda) = \frac{\left\langle \frac{H\Psi_T^{\text{lin}}}{\Psi_T^{\text{lin}}}\left(\frac{\Psi_T^{\text{lin}}}{\Psi_0}\right)^2 \right\rangle_{\Psi_0^2}}{\left\langle \left(\frac{\Psi_T^{\text{lin}}}{\Psi_0}\right)^2 \right\rangle_{\Psi_0^2}}. \tag{17}$$

IV. EVALUATING DERIVATIVES OF THE MULTI-JASTROW TRIAL WAVEFUNCTION

The quantities to be calculated at each Monte Carlo step of a VMC or fixed-node diffusion Monte Carlo (FN-DMC) simulation (see, e.g., Ref. 13) are the values of the trial wavefunction, ψ_T, its first derivatives, $\nabla \psi_T$, and Laplacian, $\nabla^2 \psi_T$, for a given electronic configuration $\mathbf{R} \equiv (\mathbf{r}_1, \ldots, \mathbf{r}_{N_{el}})$ (or a set of configurations when a population of walkers is considered). When optimizing the trial wavefunction as described in Sec. III the derivatives of ψ_T with respect to each parameter, $\partial \psi_T / \partial c_p$, and the Laplacian of these derivatives, $\nabla^2(\partial \psi_T / \partial c_p)$, must also be computed. Note that the quantities to be evaluated being linear in the products of determinants we therefore need to consider their calculation only for one elementary product of determinants, the generalization to an

arbitrary number of determinants corresponding to a simple sum. To facilitate the derivations to follow, let us introduce some notations,

$$\Psi_{ij}^{\sigma} \equiv e^{J_i(\mathbf{r}_j|\mathbf{r}_k \neq \mathbf{r}_j)}\phi_i(\mathbf{r}_j),$$

where $\sigma = \uparrow, \downarrow$. The matrices Ψ_{ij}^{σ} will be called Slater matrices and their determinants will be denoted as

$$\Psi^{\sigma} \equiv \det(\Psi_{ij}^{\sigma}).$$

For the sake of clarity we will suppose that the occupied up and down orbitals are the same ("restricted-type" calculations but using different orbitals is not a problem) so that the generic component whose various derivatives are to be computed can be written as follows:

$$\Psi = \Psi^{\uparrow}\Psi^{\downarrow}. \tag{18}$$

Before explicating the derivatives, let us present an elementary but fundamental formula we will use in the following to write down most of the expressions. Let M be a finite inversible matrix and δM a general matrix (in practice, δM will be a "small" perturbation of M resulting from the change of one or two columns or lines at most) it is trivial to verify that

$$\det(M + \delta M) = \det(M)\det(Id + M^{-1}\delta M), \tag{19}$$

where Id is the identity matrix.

A. Evaluation of Ψ

The computation of Ψ is done as usual way by using standard linear algebra techniques to compute the determinant of the matrices Ψ_{ij}^{σ}.

B. Evaluation of $\nabla \Psi$

Let us note

$$\partial_{i,l} \equiv \frac{\partial}{\partial x_i^l},$$

where i denotes the electron index running from $i=1$ to N_{σ}, and l is the space index, $x^1 \equiv x$, $x^2 \equiv y$, $x^3 \equiv z$. Using these notations we have

$$\partial_{i,l}\Psi = \sum_{k=1}^{N_{\uparrow}} \begin{vmatrix} \Psi_{11}^{\uparrow} & \cdots & \Psi_{1k-1}^{\uparrow} & \partial_{i,l}\Psi_{1k}^{\uparrow} & \Psi_{1k+1}^{\uparrow} & \cdots & \Psi_{1N_{\uparrow}}^{\uparrow} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \Psi_{j1}^{\uparrow} & \cdots & \Psi_{jk-1}^{\uparrow} & \partial_{i,l}\Psi_{jk}^{\uparrow} & \Psi_{jk+1}^{\uparrow} & \cdots & \Psi_{jN_{\uparrow}}^{\uparrow} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \Psi_{N_{\uparrow}1}^{\uparrow} & \cdots & \Psi_{N_{\uparrow}k-1}^{\uparrow} & \partial_{i,l}\Psi_{N_{\uparrow}k}^{\uparrow} & \Psi_{N_{\uparrow}k+1}^{\uparrow} & \cdots & \Psi_{N_{\uparrow}N_{\uparrow}}^{\uparrow} \end{vmatrix} \Psi^{\downarrow}$$

$$+ \Psi^{\uparrow}\sum_{k=1}^{N_{\downarrow}} \begin{vmatrix} \Psi_{11}^{\downarrow} & \cdots & \Psi_{1k-1}^{\downarrow} & \partial_{i,l}\Psi_{1k}^{\downarrow} & \Psi_{1k+1}^{\downarrow} & \cdots & \Psi_{1N_{\downarrow}}^{\downarrow} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \Psi_{j1}^{\downarrow} & \cdots & \Psi_{jk-1}^{\downarrow} & \partial_{i,l}\Psi_{jk}^{\downarrow} & \Psi_{jk+1}^{\downarrow} & \cdots & \Psi_{jN_{\downarrow}}^{\downarrow} \\ \vdots & & \vdots & \vdots & \vdots & & \vdots \\ \Psi_{N_{\downarrow}1}^{\downarrow} & \cdots & \Psi_{N_{\downarrow}k-1}^{\downarrow} & \partial_{i,l}\Psi_{N_{\downarrow}k}^{\downarrow} & \Psi_{N_{\downarrow}k+1}^{\downarrow} & \cdots & \Psi_{N_{\downarrow}N_{\downarrow}}^{\downarrow} \end{vmatrix}. \tag{20}$$

To evaluate the determinants the following decomposition of the matrices corresponding to the determinants ↑ and ↓ are considered:

$$
M^\sigma + \delta M_k^\sigma = \begin{pmatrix}
\Psi_{11}^\sigma & \cdots & \Psi_{1k-1}^\sigma & \partial_{i,l}\Psi_{1k}^\sigma & \Psi_{1k+1}^\sigma & \cdots & \Psi_{1N_\sigma}^\sigma \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
\Psi_{j1}^\sigma & \cdots & \Psi_{jk-1}^\sigma & \partial_{i,l}\Psi_{jk}^\sigma & \Psi_{jk+1}^\sigma & \cdots & \Psi_{jN_\sigma}^\sigma \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
\Psi_{N_\sigma 1}^\sigma & \cdots & \Psi_{N_\sigma k-1}^\sigma & \partial_{i,l}\Psi_{N_\sigma k}^\sigma & \Psi_{N_\sigma k+1}^\sigma & \cdots & \Psi_{N_\sigma N_\sigma}^\sigma
\end{pmatrix},
\tag{21}
$$

where M^σ is the Slater matrix, Ψ_{ij}^σ, and δM_k^σ is a "perturbing" matrix having only the kth column different from zero,

$$
(\delta M_k^\sigma)_j \equiv -\Psi_{jk}^\sigma + \partial_{i,l}\Psi_{jk}^\sigma, \quad j = 1, N_\sigma.
\tag{22}
$$

The matrix $Id + [M^\sigma]^{-1}\delta M_k^\sigma$ has thus the following structure:

$$
Id + [M^\sigma]^{-1}\delta M_k^\sigma = \begin{pmatrix}
1 & 0 & \cdots & (P_k^\sigma)_1 & \cdots & 0 & 0 \\
0 & 1 & \cdots & (P_k^\sigma)_2 & \cdots & 0 & 0 \\
\vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\
\vdots & \vdots & \vdots & 1+(P_k^\sigma)_k & & \vdots & \vdots \\
\vdots & \vdots & \vdots & \vdots & & \vdots & \vdots \\
0 & 0 & \cdots & (P_k^\sigma)_{N_\sigma-1} & \cdots & 1 & 0 \\
0 & 0 & \cdots & (P_k^\sigma)_{N_\sigma} & \cdots & 0 & 1
\end{pmatrix},
\tag{23}
$$

where the vector (P_k^σ) is defined as

$$
(P_k^\sigma)_m = \sum_{j=1}^{N_\sigma} [\Psi^\sigma]_{mj}^{-1}(\delta M_k^\sigma)_j, \quad m = 1, N_\sigma.
\tag{24}
$$

Now, applying formula Eq. (19) we get

$$
\det(M^\sigma + \delta M_k^\sigma) = \Psi^\sigma[1 + (P_k^\sigma)_k]
\tag{25}
$$

and, finally, the expression of the derivative of Ψ is given by

$$
\partial_{i,l}\Psi = \Psi\left\{ \sum_{k=1}^{N_\uparrow} [1 + (P_k^\uparrow)_k] + \sum_{k=1}^{N_\downarrow} [1 + (P_k^\downarrow)_k] \right\}.
\tag{26}
$$

As seen this expression requires the computation of the Slater matrix elements and their derivatives, and also the inverse of the Slater matrices which are computed in standard implementation of QMC algorithms (see, e.g., Ref. 13).

C. Evaluation of $\nabla^2\Psi$

The second derivatives of Ψ can be written as

$$
\partial_{i,l}^2\Psi = (\partial_{i,l}^2\Psi^\uparrow)\Psi^\downarrow + \Psi^\uparrow\partial_{i,l}^2\Psi^\downarrow + 2\partial_{i,l}\Psi^\uparrow\partial_{i,l}\Psi^\downarrow.
\tag{27}
$$

The first derivatives, $\partial_{i,l}\Psi^\sigma$, are computed as presented in the preceding section. The new terms, $\partial_{i,l}^2\Psi^\sigma$, can be written as

$$
\partial_{i,l}^2\Psi^\sigma = \sum_{k=1}^{N_\sigma}\sum_{m=1}^{N_\sigma} \begin{vmatrix}
\Psi_{11}^\sigma & \cdots & \partial_{i,l}\Psi_{1k}^\sigma & \cdots & \partial_{i,l}\Psi_{1m}^\sigma & \cdots & \Psi_{1N_\sigma}^\sigma \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
\Psi_{j1}^\sigma & \cdots & \partial_{i,l}\Psi_{jk}^\sigma & \cdots & \partial_{i,l}\Psi_{jm}^\sigma & \cdots & \Psi_{jN_\sigma}^\sigma \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
\Psi_{N_\sigma 1}^\sigma & \cdots & \partial_{i,l}\Psi_{N_\sigma k}^\sigma & \cdots & \partial_{i,l}\Psi_{N_\sigma m}^\sigma & \cdots & \Psi_{N_\sigma N_\sigma}^\sigma
\end{vmatrix},
\tag{28}
$$

where the derivatives appear only in columns k and m. Let us distinguish the cases corresponding to $k=m$ and $k \neq m$.

- Case $k=m$: In this case where the column $k=m$ is given by the vector $\partial_{i,l}^2\Psi_{jk}^\sigma$ with $j=1$, N_σ, we are in a situation absolutely similar to what we had before with the first derivative. The perturbing matrix δM^σ to consider is the matrix having as unique nonzero column the following vector

$$
(\delta M_k^\sigma)_j \equiv -\Psi_{jk}^\sigma + \partial_{i,l}^2\Psi_{jk}^\sigma, \quad j = 1, N_\sigma,
\tag{29}
$$

and the value of the second derivative is given by

$$
\partial_{i,l}^2\Psi^\sigma = \Psi^\sigma\sum_{k=1}^{N_\sigma} [1 + (P_k^\sigma)_k],
\tag{30}
$$

where

$$
(P_k^\sigma)_k = \sum_{j=1}^{N_\sigma} [\Psi^\sigma]_{kj}^{-1}(\delta M_k^\sigma)_j.
\tag{31}
$$

- Case $k \neq m$: This case leads to a new situation where the perturbing matrix δM_{km}^σ consists now of two nonvanishing columns k and m represented by the vectors $-\Psi_{jk}^\sigma + \partial_{i,l}\Psi_{jk}^\sigma$ and $-\Psi_{jm}^\sigma + \partial_{i,l}\Psi_{jm}^\sigma$. Defining as before the vectors P_k^σ and P_m^σ as follows:

$$
(P_k^\sigma)_n = \sum_{j=1}^{N_\sigma} [\Psi^\sigma]_{nj}^{-1}(\delta M_k^\sigma)_j,
\tag{32}
$$

$$(P_m^\sigma)_n = \sum_{j=1}^{N_\sigma} [\Psi^\sigma]_{nj}^{-1} (\delta M_m^\sigma)_j .$$

The following structure for the matrix $Id + [M^\sigma]^{-1} \delta M_{km}^\sigma$ is obtained:

$$Id + [M^\sigma]^{-1} \delta M_{km}^\sigma$$

$$= \begin{pmatrix}
1 & \cdots & (P_k^\sigma)_1 & \cdots & 0 & \cdots & (P_m^\sigma)_1 & \cdots & 0 \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
0 & \cdots & 1+(P_k^\sigma)_k & \cdots & 0 & \cdots & (P_m^\sigma)_k & \cdots & 0 \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
0 & \cdots & (P_k^\sigma)_m & \cdots & 0 & \cdots & 1+(P_m^\sigma)_m & \cdots & 0 \\
\vdots & & \vdots & & \vdots & & \vdots & & \vdots \\
0 & \cdots & (P_k^\sigma)_{N_\sigma} & \cdots & 0 & \cdots & (P_m^\sigma)_{N_\sigma} & \cdots & 1
\end{pmatrix} .$$

$$(33)$$

The determinant of this matrix is given by $[(1+P_k^\sigma)_k] \times [(1+P_m^\sigma)_m] - (P_m^\sigma)_k (P_k^\sigma)_m$ and, finally, the value of the second derivative can be written as

$$\partial_{i,l}^2 \Psi^\sigma = \sum_{k=1}^{N_\sigma} \sum_{m=1}^{N_\sigma} \Psi^\sigma \{ [1+(P_k^\sigma)_k][1+(P_m^\sigma)_m] - (P_m^\sigma)_k (P_k^\sigma)_m \} .$$

$$(34)$$

D. Derivative of Ψ with respect to a local Jastrow parameter

In this section we evaluate the derivative of Ψ with respect to a parameter c_p appearing in the expression of a Jastrow term, J_i. Here, we shall only consider the case where this parameter belongs only to a single Jastrow (no common parameter for distinct Jastrows). The generalization to the case where the parameter can appear in various Jastrows does not bring special difficulties. Denoting ∂_{c_p} the derivative with respect to c_p we have

$$\partial_{c_p} \Psi = \sum_{k=1}^{N_\uparrow} \begin{vmatrix}
\Psi_{11}^\uparrow & \cdots & \Psi_{1j}^\uparrow & \cdots & \Psi_{1N_\uparrow}^\uparrow \\
\vdots & & \vdots & & \vdots \\
\partial_{c_p}\Psi_{k1}^\uparrow & \cdots & \partial_{c_p}\Psi_{kj}^\uparrow & \cdots & \partial_{c_p}\Psi_{kN_\uparrow}^\uparrow \\
\vdots & & \vdots & & \vdots \\
\Psi_{N_\uparrow 1}^\uparrow & \cdots & \Psi_{N_\uparrow j}^\uparrow & \cdots & \Psi_{N_\uparrow N_\uparrow}^\uparrow
\end{vmatrix} \Psi^\downarrow$$

$$+ \Psi^\uparrow \sum_{k=1}^{N_\downarrow} \begin{vmatrix}
\Psi_{11}^\downarrow & \cdots & \Psi_{1j}^\downarrow & \cdots & \Psi_{1N_\downarrow}^\downarrow \\
\vdots & & \vdots & & \vdots \\
\partial_{c_p}\Psi_{k1}^\downarrow & \cdots & \partial_{c_p}\Psi_{kj}^\downarrow & \cdots & \partial_{c_i}\Psi_{kN_\downarrow}^\downarrow \\
\vdots & & \vdots & & \vdots \\
\Psi_{N_\downarrow 1}^\downarrow & \cdots & \Psi_{N_\downarrow j}^\downarrow & \cdots & \Psi_{N_\downarrow N_\downarrow}^\downarrow
\end{vmatrix} .$$

$$(35)$$

As seen, this expression is similar to that obtained for the gradient of Ψ, Eq. (20), except that the derivatives occur on one line and not on a column. It is thus simple to verify that the same derivation applies, we get

$$\partial_{c_p} \Psi = \Psi \left\{ \sum_{k=1}^{N_\uparrow} [1 + (P_k^\uparrow)_k] + \sum_{k=1}^{N_\downarrow} [1 + (P_k^\downarrow)_k] \right\} ,$$

$$(36)$$

with

$$(P_k^\sigma)_k = \sum_{j=1}^{N_\sigma} [\Psi^\sigma]_{kj}^{-1} (\delta M_k^\sigma)_j$$

$$(37)$$

and

$$(\delta M_k^\sigma)_j = -\Psi_{jk}^\sigma + \partial_{c_p}\Psi_{jk}^\sigma, \quad j = 1, N_\sigma.$$

$$(38)$$

E. Derivative of $\nabla^2\Psi$ with respect to a local Jastrow parameter

The derivative of $\nabla^2(\Psi^\uparrow \Psi^\downarrow)$ with respect to a Jastrow's parameter c_p can be decomposed in a series of elementary terms. Most of these terms have been already derived in the previous sections, namely, Ψ^σ, $\nabla\Psi^\sigma$, $\nabla^2\Psi^\sigma$, and $\partial\Psi^\sigma / \partial c_p$. The remaining terms to be computed are the gradient and Laplacian of the derivative of Ψ^σ, namely, $\nabla(\partial\Psi^\sigma / \partial c_p)$ and $\nabla^2(\partial\Psi^\sigma / \partial c_p)$.

1. Evaluation of $\nabla(\partial\Psi^\sigma / \partial c_p)$

Using the notations introduced above the gradient of Ψ^σ is written as

$$\partial_{i,l}\Psi^\sigma = \sum_{k=1}^{N_\sigma} \begin{vmatrix}
\Psi_{11}^\sigma & \cdots & \Psi_{1k-1}^\sigma & \partial_{i,l}\Psi_{1k}^\sigma & \Psi_{1k+1}^\sigma & \cdots & \Psi_{1N_\sigma}^\sigma \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
\Psi_{j1}^\sigma & \cdots & \Psi_{jk-1}^\sigma & \partial_{i,l}\Psi_{jk}^\sigma & \Psi_{jk+1}^\sigma & \cdots & \Psi_{jN_\sigma}^\sigma \\
\vdots & & \vdots & \vdots & \vdots & & \vdots \\
\Psi_{N_\sigma 1}^\sigma & \cdots & \Psi_{N_\sigma k-1}^\sigma & \partial_{i,l}\Psi_{N_\sigma k}^\sigma & \Psi_{N_\sigma k+1}^\sigma & \cdots & \Psi_{N_\sigma N_\sigma}^\sigma
\end{vmatrix}$$

$$(39)$$

and the derivative with respect to Jastrow's parameter c_p as

$$\partial_{c_p}\partial_{i,l}\Psi^\sigma = \sum_{j=1}^{N_\sigma}\sum_{k=1}^{N_\sigma}\begin{vmatrix} \Psi_{11}^\sigma & \cdots & \partial_{i,l}\Psi_{1k}^\sigma & \cdots & \Psi_{1N_\sigma}^\sigma \\ \vdots & & \vdots & & \\ \partial_{c_p}\Psi_{j1}^\sigma & \cdots & \partial_{c_p}\partial_{i,l}\Psi_{jk}^\sigma & \cdots & \partial_{c_p}\Psi_{jN_\sigma}^\sigma \\ \vdots & & \vdots & & \\ \Psi_{N_\sigma 1}^\sigma & \cdots & \partial_{i,l}\Psi_{jk}^\sigma & \cdots & \Psi_{N_\sigma N_\sigma}^\sigma \end{vmatrix}. \tag{40}$$

To compute the new determinants appearing in this expression, the fundamental formula, Eq. (19), is used with

$$M^\sigma = \begin{pmatrix} \Psi_{11}^\sigma & \cdots & \Psi_{1k}^\sigma & \cdots & \Psi_{1N_\sigma}^\sigma \\ \vdots & & \vdots & & \\ \partial_{c_p}\Psi_{j1}^\sigma & \cdots & \partial_{c_p}\Psi_{jk}^\sigma & \cdots & \Psi_{jN_\sigma}^\sigma \\ \vdots & & \vdots & & \\ \Psi_{N_\sigma 1}^\sigma & \cdots & \Psi_{jk}^\sigma & \cdots & \Psi_{N_\sigma N_\sigma}^\sigma \end{pmatrix} \tag{41}$$

and

$$\delta M^\sigma = \begin{pmatrix} 0 & \cdots & -\Psi_{1k}^\sigma + \partial_{i,l}\Psi_{1k}^\sigma & \cdots & 0 \\ \vdots & & \vdots & & \\ 0 & \cdots & -\partial_{c_p}\Psi_{jk}^\sigma + \partial_{c_p}\partial_{i,l}\Psi_{jk}^\sigma & \cdots & 0 \\ \vdots & & \vdots & & \\ 0 & \cdots & -\Psi_{jk}^\sigma + \partial_{i,l}\Psi_{jk}^\sigma & \cdots & 0 \end{pmatrix}. \tag{42}$$

How to calculate the determinant of M has already been presented. Now, to employ Eq. (19) we also need to evaluate the inverse of M. To avoid costly N_σ^3 evaluations using standard methods we propose to employ the Sherman–Morrison formula.[14] Let us denote A as an inversible matrix and u and v two vectors, this formula is written as

$$(A + u\otimes v)^{-1} = A^{-1} - \frac{z\otimes w}{1+\lambda}, \tag{43}$$

where

$$\lambda \equiv (v, A^{-1}u), \quad z\equiv A^{-1}u, \quad w\equiv (A^{-1})^T v$$

and

$$[u\otimes v]_{ij} \equiv u_i v_j.$$

Here, to calculate the inverse of matrix M we apply the Sherman–Morrison formula with the two vectors

$$u_i = -\Psi_{li}^\sigma + \frac{\partial\Psi_{li}^\sigma}{c_l} \tag{44}$$

and

$$v_i = \delta_{il}. \tag{45}$$

2. Evaluation of $\nabla_i^2(\partial\Psi^\sigma/\partial c_p)$

To compute $\nabla_i^2(\partial\Psi^\sigma/\partial c_p)$ it is easy to check that we need to evaluate the two types of determinant

TABLE I. VMC and FN-DMC total ground-state energies for the ^3P oxygen atom using different trial wavefunctions and nodal structures. All Jastrows (common or individual) are fully optimized.

Method	Energy (a.u.)	σ^2 (energy)	Correlation energy
HF[a]	−74.8094	⋯	0%
"Exact" nonrelativistic[b]	−75.0673	⋯	100%
VMC			
HF (no Jastrow of any type)	−74.808(2)	15.9(3)	0.% ± 0.8%
HF+common Jastrow	−75.025(1)	1.178(7)	83.6% ± 0.4%
HF+core- and val-Jastrows (J_{1s}, J_{2s2p})	−75.0275(4)	1.133(8)	84.6% ± 0.2%
FN-DMC			
HF nodes	−75.052(1)		94.1% ± 0.4%
Multi-Jastrow nodes	−75.052(1)		94.1% ± 0.4%

[a]Reference 16.
[b]Reference 20.

$$\begin{vmatrix} \Psi_{11}^\sigma & \cdots & \Delta_i(\Psi_{1k}^\sigma) & \cdots & \Psi_{1N_\sigma}^\sigma \\ \vdots & & \vdots & & \\ \partial_{c_p}\Psi_{j1}^\sigma & \cdots & \partial_{c_p}\Delta_i(\Psi_{jk}^\sigma) & \cdots & \partial_{c_p}\Psi_{jN_\sigma}^\sigma \\ \vdots & & \vdots & & \\ \Psi_{N_\sigma 1}^\sigma & \cdots & \Delta_i(\Psi_{jk}^\sigma) & \cdots & \Psi_{N_\sigma N_\sigma}^\sigma \end{vmatrix} \tag{46}$$

and

$$\begin{vmatrix} \Psi_{11}^\sigma & \cdots & \partial_{i,l}\Psi_{1k}^\sigma & \cdots & \Psi_{1N_\sigma}^\sigma \\ \vdots & & \vdots & & \vdots \\ \partial_{c_p}\Psi_{j1}^\sigma & \cdots & \partial_{c_p}\partial_{i,l}\Psi_{jk}^\sigma & \cdots & \partial_{c_p}\Psi_{jN_\sigma}^\sigma \\ \vdots & & \vdots & & \vdots \\ \Psi_{N_\sigma 1}^\sigma & \cdots & \partial_{i,l}\Psi_{jk}^\sigma & \cdots & \Psi_{N_\sigma N_\sigma}^\sigma \end{vmatrix} \tag{47}$$

and we make use here also of our fundamental formula combined with the Sherman–Morrison formula.

V. RESULTS

A. Atoms

As a first application of the use of the multi-Jastrow wavefunction we have considered atomic systems. Numerous studies on atoms using QMC techniques have been presented in the last 20 years (see, for example, Ref. 15). Here, we shall consider three systems of increasing size, namely, the ^3P ground-state oxygen atom ($1s^22s^22p^4$, 8 electrons), the ^3P ground-state sulfur atom ($[Ne]3s^23p^4$, 18 electrons), and the ^2S ground-state copper atom ($[Ar]3d^{10}4s^1$, 29 electrons).

1. The oxygen atom

Table I presents our results for the oxygen atom. The total energies obtained at the VMC and FN-DMC levels with different trial wavefunctions are presented. Two different criteria to quantify the quality of the trial wavefunctions are employed: the variance of the energy, denoted as $\sigma^2(E)$, and

the amount of correlation energy recovered. The variance of the energy is defined as the variance of the local energy during the simulation

$$\sigma^2(E) \equiv \langle E_L^2 \rangle - \langle E_L \rangle^2, \tag{48}$$

where the symbol $\langle \cdots \rangle$ denotes the Monte Carlo average (here, either VMC or FN-DMC) and E_L is the local energy, the basic quantity of QMC methods, defined as $E_L \equiv H\Psi_T/\Psi_T$. Note that the variance of the energy is a natural quantity to introduce in a QMC framework since the statistical error is directly proportional to its square root. In particular, the lower the variance of the local energy is the "closer" the trial wavefunction is from the exact wavefunction. The second criterion is standard in quantum chemistry: The correlation energy defined as the difference between the exact nonrelativistic Born–Oppenheimer energy and the HF energy obtained with an infinite one-particle basis set (complete basis set limit or "HF limit").

At the VMC level, three different trial wavefunctions have been employed. The first one is a standard HF wavefunction built from the atomic orbitals expressed over Slater-type orbitals as given by Bunge et al.[16] As it should be, the VMC result recovers within statistical uncertainties the exact variational energy. The second trial wavefunction has the standard form consisting of a Slater determinant multiplied by a Jastrow factor, Eq. (6). All 19 parameters of the Jastrow term have been optimized, while the atomic orbitals have been kept unchanged. The optimization leads to a gain of 83.6% of the correlation energy, a typical result for light atoms (see, e.g., Ref. 15 where Schmidt and Moskowitz recovered about 80% of the correlation energy for the oxygen atom). Finally, the last trial wavefunction is our multi-Jastrow wavefunction. Using a trial-and-error approach it has been found that only two different atomic Jastrows were sufficient to represent the atom. The first one attached to the innermost $1s$ orbital is called the core-Jastrow while the second one attached to the $2s$ and $2p$ atomic orbitals will be referred to as the valence-Jastrow. After full optimization of the two local Jastrow terms (12 parameters each) the VMC energy obtained is only slightly lower than the energy resulting from the use of a common Jastrow. This result is of course disappointing but illustrates that at least for the oxygen atom, making the difference between the core and valence regions for the Jastrow part does not lead to significant improvement compared to the usual case based on the use of a global Jastrow.

This conclusion also extends to the FN-DMC results. Indeed, as seen in Table I the FN-DMC results obtained with HF or multi-Jastrow nodes are identical within statistical uncertainties. This result seems to indicate that for this simple atomic system the nodal properties of the trial wavefunction are not improved when introducing different local Jastrow terms.

2. The sulfur atom

The VMC and DMC results for the sulfur atom are presented in Table II. Although the number of electrons is greater than in the previous case and although the electronic density is more contrasted (a point which should, in prin-

ciple, favor the multi-Jastrow form), the results obtained are quite similar. The amount of correlation energy recovered in the standard case with the common Jastrow, $80.2\% \pm 0.16\%$, is almost identical to what we have obtained with a three-Jastrow form (a different Jastrow per atomic shell), namely, $81.1\% \pm 0.1\%$. In addition, as in the oxygen case, the FN-DMC energies obtained using the two types of nodal structure associated with the HF and multi-Jastrow forms are almost identical.

3. The copper atom

The case of the copper atom is much more interesting since for a transition metal atom with many d electrons occupying a small region of space, it is expected that electronic correlations may differ drastically from one shell to the other. The VMC and FN-DMC results obtained are presented in Table III. At the VMC level with the standard Jastrow–Slater form for the trial wavefunction about 61% of the correlation energy is recovered. In contrast, our multi-Jastrow form allows us to recover a better value, namely, about 75% of the correlation energy. To get this very good result, three different local Jastrow terms have been considered: one for the $1s$ core orbital, another for the $2s2p3s3p$ orbitals, and the last one for the $3d$ and $4s$ orbitals. It should be emphasized that this result is very satisfactory since it gives strong support for the validity of distinguishing electronic correlations among the different shells of a structured system such as the copper atom by using different Jastrow factors. Note also that the total variational energy obtained here for the copper atom, $E=-1640.173(7)$ could probably be further improved by using more sophisticated forms for our individual Jastrows, Eq. (10), and also by fully optimizing the atomic orbitals of the atom. However, our objective here was not to get the most accurate value of the total energy but rather to show that by using a multi-Jastrow form instead of a standard one, results can be significantly improved (to allow a fair comparison, a common set of atomic orbitals and a common functional form for the Jastrow terms have been used).

Now, regarding the FN results for the total DMC energies the situation is found to be similar to what has been obtained for the oxygen and sulfur atoms. Indeed, as seen in Table III no improvement of the nodes is observed when going from the HF to the new multi-Jastrow wavefunctions, despite the clear improvement of the variational energy.

General conclusion for atoms. A first remark is that the results presented here for the first-row (oxygen) and second-row (sulfur) atoms do not show any significant improvement when going from the standard form (with a global Jastrow) to the multi-Jastrow form proposed in this work. Note that additional results for other light elements have been made and this negative conclusion has systematically been found valid. In contrast, in the case of the copper atom having a maximally filled $3d$-shell it has been possible to obtain a variational energy significantly better than in the standard case, thus illustrating the potential interest of the multi-Jastrow wavefunction. However, for all atoms no improvement regarding the FN results has been found. A possible interpretation of the good result obtained at the variational level for the copper atom could be as follows. In the case of

TABLE II. VMC and FN-DMC total ground-state energies for the ^3P sulfur atom using different trial wavefunctions and nodal structures. All Jastrows (common or individual) are fully optimized.

Method	Energy (a.u.)	σ^2 (energy)	Correlation energy
HF[a]	−397.5049	⋯	0%
Exact nonrelativistic[b]	−398.111	⋯	100%
	VMC		
HF (no Jastrow of any type)	−397.505(1)	97.0(4)	0.% ± 0.16%
HF+common Jastrow	−397.991(1)	14.9(1)	80.2% ± 0.16%
HF+3 Jastrows ($J_{n=1}, J_{n=2}, J_{n=3}$)	−397.9964(6)	9.29(2)	81.1% ± 0.1%
	FN-DMC		
HF nodes	−398.071(1)		93.4% ± 0.16%
Multi-Jastrow nodes	−398.068(2)		92.9% ± 0.32%

[a]Reference 16.
[b]Reference 20.

TABLE III. VMC and FN-DMC total ground-state energies for the ^2S copper atom using different trial wavefunctions and nodal structures. All Jastrows (common or individual) are fully optimized.

Method	Energy (a.u.)	σ^2 (energy)	Correlation energy
HF[a]	−1638.9637	⋯	0%
Exact nonrelativistic[b]	−1640.5677	⋯	100%
	VMC		
HF (no Jastrow of any type)	−1638.970(6)	423(7)	0.% ± 0.8%
HF+common Jastrow	−1639.936(5)	75(1)	60.6% ± 0.3%
HF+3 Jastrows ($J_{1s}, J_{2s2p3s3p}, J_{3d4s}$)	−1640.1725(74)	131(3)	75.4% ± 0.5%
	FN-DMC		
HF nodes	−1640.401(9)		89.6% ± 0.6%
Multi-Jastrow nodes	−1640.407(8)		90.0% ± 0.5%

[a]Reference 16.
[b]Reference 20.

the light atoms of the first- and second-rows it seems reasonable to think that a common Jastrow having enough flexibility to give a good description of the electron-electron, electron-nucleus, and electron-electron-nucleus interactions could also describe correctly the differential effects corresponding to the $n=1$, $n=2$, and $n=3$ shells with s and p orbitals. In contrast, the correct description of the behavior of the many d-electron occupying a small region of space is very likely qualitatively different of what happens in the gentler and less crowded $2s, 2p, 3s, 3p$ orbitals. Accordingly, to be able to separate the two types of local correlation effects are certainly a major advantage of our new proposed form. Now, regarding the FN data the results obtained are much more puzzling. The presence of a spherical symmetry plus the use of fixed molecular orbitals (the 3D-nodes of the orbitals are not changed) might impose relatively tight constraints on the nodal surfaces and prevent nodal deformations. However, this point is not clear and is left for further investigation.

B. The FH molecule

As a first example of a molecular application, the calculation of the potential energy curve of the FH molecule has been considered. In particular, the three basic spectroscopic constants describing the overall shape of the curve—the equilibrium distance r_e, the harmonic frequency ω_e, and the dissociation energy D_0—are computed.

1. Taking care of the static correlation effects: Use of a minimal CASSCF wavefunction

Before introducing our molecular Jastrows it is important to impose to the initial *ab initio* wavefunction on which the new multi-Jastrow form is built to have the correct dissociation property. Indeed, to get a clear assessment of the quality of the proposed form, the various types of electron correlations at work must absolutely be properly distinguished. To be more precise, the quality of the multi-Jastrow nodal pattern is expected to be intimately related to the abil-

ity of such a wavefunction to correctly describe both *nondynamical* (*static*) and *dynamical* correlation effects. By introducing our local Jastrow factors we are essentially improving the dynamical correlation effects and very little the static ones related to near-degeneracy effects. Accordingly, introducing the major part of these latter effects via an appropriate CASSCF representation is an essential prerequisite.

Here, the minimal form leading to the correct dissociation of the molecule into its neutral atomic fragments has been considered. More precisely, the Slater part is given by the following two-determinant form:

$$\Psi_{T\,\mathrm{det}} = [\sigma_{1s}^2 \sigma_{2s}^2 \pi_x^2 \pi_y^2]\sigma^2 + [\sigma_{1s}^2 \sigma_{2s}^2 \pi_x^2 \pi_y^2]\sigma^{*2}, \tag{49}$$

where the molecular orbitals σ_{1s}, σ_{2s}, π_x, and π_y have a strong atomic character corresponding to the fluor atom and (σ, σ^*) are the bonding/antibonding valence orbitals built for the most part from the $1s^H$ and $2p_z^F$ atomic orbitals.

In Fig. 1 the various CASSCF curves obtained with this trial wavefunction and basis sets of increasing size are presented [(correlation-consistent basis sets of Dunning and co-workers, cc-pVnZ with $n=2$ to $n=5$ (Ref. 17)] and compared to the estimate exact nonrelativistic curve.[18] As seen, the size of the basis set has a strong impact on the overall shape of the energy curves. In our example, results seem to be converged only when basis sets corresponding to $n=4$ and beyond are considered. To have a more quantitative assessment of the quality of these curves, the spectroscopic constants corresponding to each curve are presented in Table IV. For completeness, we also give the SCF results corresponding to the monodeterminantal configuration $[\sigma_{1s}^2 \sigma_{2s}^2 \pi_x^2 \pi_y^2]\sigma^2$.

As well-known, SCF results lead to too short equilibrium distances and to too large harmonic frequencies. Here, for large basis sets r_e is too small by about 2% and ω_e is too large by about 16%. The reason for that is related to the wrong curvature of the SCF potential energy curve due to the incorrect dissociation of the SCF representation (energy curve too "stiff" at large distances because of the undesirable high-energy ionic components of the SCF wavefunction). Using now a minimal CASSCF wavefunction, as in Eq. (49),

this problem is easily avoided and both equilibrium distance and harmonic frequency are greatly improved with an error of 0.25% (2% in the SCF case) and 1% (16% in the SCF case), respectively. Regarding dissociation energies, the CASSCF results as a function of the basis set are found to converge to about 115 kcal/mol, a quantity much too small by about 26 kcal/mol. Note that the SCF dissociation energy cannot be defined due to the wrong dissociation property.

2. FN-DMC results

a. Pure CASSCF nodes.   Let us begin by discussing the FN-DMC results obtained with trial wavefunctions having the nodal pattern of the CASSCF wavefunctions, Eq. (49). Such FN-DMC calculations can be considered as standard since they are based on the usual strategy consisting in using trial wavefunctions of the Jastrow–Slater-type with optimization of the Jastrow part only. Note that for such a small molecular system, it would also be possible to perform a full optimization of the trial wavefunction (molecular orbitals, determinantal weights, and Jastrow parameters—see, e.g., Refs. 11 and 5). However, such a strategy being of practical interest only for small systems it will not be considered here. Let us look at the impact of the FN error on the FH molecular system. To see the effect of the basis set on the CASSCF nodal pattern we present in Fig. 2 the various curves obtained using the different basis sets introduced above. The corresponding spectroscopic constants are presented in Table V. Note that the extraction of such constants from noisy (Monte Carlo) potential energy curves requires some care. Here, to calculate the equilibrium distance R_e and the harmonic frequency ω_e the potential energy curves have been systematically fitted using a five-parameter generalized Morse potential form as follows:

$$E(r) = p_1 + (p_2 - p_1)e^{-p_3 x} + p_5(e^{-2p_3 x} - e^{-p_3 x}), \quad (50)$$

with

$$x = \frac{r - p_4}{p_4}. \quad (51)$$

In this formula each parameter p_i is directly related to one of the main features of the energy curve. In this way, the fit is "physically" meaningful and is neither redundant nor insufficient. The parameter p_1 is directly related to $E(r \to +\infty)$, p_2 to $E(r_e)$, p_3 to the curvature at the minimum, p_4 to r_e, and p_5 to the way the curve goes up at small r (or, equivalently, to the asymmetry of the curve around the minimum). To calculate the dissociation energy, ΔE, the difference of total energies corresponding to the equilibrium distance extracted from the fitted curve and the total energy at $R=20$ a.u. has been evaluated.

A first remark concerning the FN energy curves for FH is that the DMC results are rather accurate. For example, at the minimum, the total correlation energy recovered with the largest basis set is about 94% (a typical result for a first-row diatomic molecule, see Ref. 19). As seen in Fig. 2 the FN curves have a significant dependence on the dimension of the
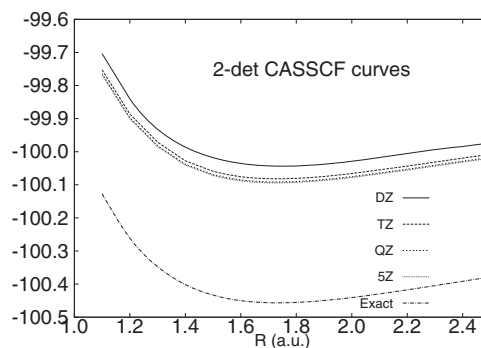

FIG. 1. Potential energy curves of the FH molecule calculated at the 2-det CASSCF level using various basis sets: cc-pVnZ, $n=2$, 3, 4, and 5. The exact nonrelativistic curve at the bottom of the figure is obtained from Ref. 18.

basis set. At the DZ level (smallest basis set), the FN error is much larger than with the other basis sets and we conclude that such a basis is not appropriate. In the neighborhood of the minimum and for the three largest basis sets (TZ, QZ, and 5Z) the dependence is much weaker and similar results are obtained. However, at the largest distances shown (R between 2 and 2.4), it is no longer the case and the nodal pattern appears not to be fully stabilized as a function of the basis set. To get a more quantitative view, let us now turn our attention to the spectroscopic constants presented in Table V. As a function of the basis set the dissociation energy is found to converge around -138 kcal/mol. This value is greatly improved with respect to the CASSCF value of -115 kcal/mol presented in Table V but is still in error by a contribution of about 3 kcal/mol. Now, the equilibrium distances and the harmonic frequencies are acceptable but are not stabilized as a function of the basis set used. This fact is a direct consequence of the lack of stabilization of the DMC energies at the largest distances as already pointed out.

b. Multi-Jastrow/CASSCF nodes.   Let us now look at the results obtained with the multi-Jastrow wavefunction. Different choices regarding to which molecular orbitals should be attributed a molecular Jastrow term are possible. Quite clearly, the innermost orbital is almost purely atomic (1s-like) and does not require a Jastrow term as far as bonding properties are considered. We have found that introduc-

TABLE IV. Spectroscopic constants of the FH molecule at the SCF and minimal CASSCF levels as a function of the basis set.

Method	R_e (a.u.)	ω_e (cm^{-1})	D_0 (kcal/mol)
SCF (cc-pVDZ basis)	1.7042	4727	⋯
SCF (cc-pVTZ basis)	1.6971	4829	⋯
SCF (cc-pVQZ basis)	1.6956	4806	⋯
SCF (cc-pV5Z basis)	1.6953	4805	⋯
2-det CASSCF (cc-pVDZ basis)	1.7412	3960	109.27
2-det CASSCF (cc-pVTZ basis)	1.7309	4083	113.88
2-det CASSCF (cc-pVQZ basis)	1.7283	4088	114.80
2-det CASSCF (cc-pV5Z basis)	1.7283	4096	115.04
Exact	1.7326	4137	141.1[a]

[a]Obtained by using the experimental estimate of the correlation energy, E_c =0.3882 a.u. (Ref. 21), the exact atomic energy values (Ref. 20), and the SCF limit for FH (Ref. 18).
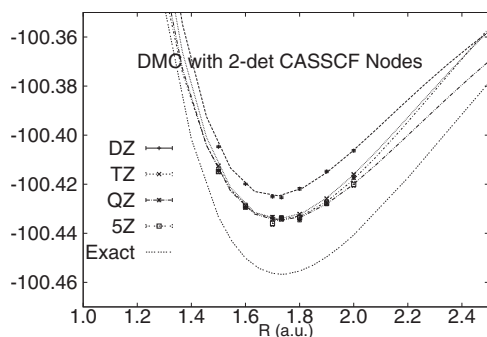
FIG. 2. Potential energy curves of the FH molecule calculated at the FN-DMC level using as trial wavefunction 2-det CASSCF wavefunctions obtained from various basis sets: cc-pVnZ, $n=2$, 3, 4, and 5. The exact non-relativistic curve at the bottom of the figure is obtained from Ref. 18.
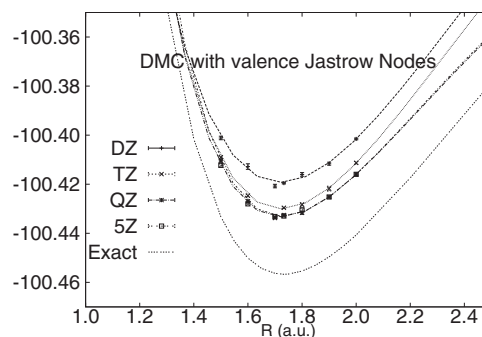
FIG. 3. Potential energy curves of the FH molecule calculated at the FN-DMC level using as trial wavefunction the multi-Jastrow wavefunction built from 2-det CASSCF wavefunction with various basis sets: cc-pVnZ, $n=2$, 3, 4, and 5. The exact nonrelativistic curve at the bottom of the figure is obtained from Ref. 18.

ing a Jastrow term for the fluor lone pair corresponding essentially to the $2s$ atomic orbital has also a negligible impact on the molecular properties (the $2s$ energy is quite low). Finally, it seems reasonable to consider *two molecular Jastrow terms*, one attached to the two lone pairs related to the atomic lone pairs $2p_x$ and $2p_y$ (z-internuclear axis) having a higher energy than in the $2s$ case, and one specific to the bonding/antibonding molecular orbitals built essentially from the $2p_z$ and $1s_H$ orbitals.

The curves obtained using various basis sets are presented in Fig. 3 and the corresponding spectroscopic values are presented in Table V. The global shape of the curves is similar to the curves obtained previously. However, some important differences can be observed. A first important difference concerns the values of the total energies obtained with the multi-Jastrow form. When compared to the standard case with CASSCF nodes, the multi-Jastrow FN energy values are found to be slightly *higher* by a quantity dependent both on the internuclear distance and on the basis set (typically, this quantity is between 0.001 and 0.005 a.u.). At first sight, such a result could be considered as surprising or disappointing. However, it should not be the case since the quality of the new multi-Jastrow wavefunction must not be judged by considering global properties such as the total

energy (whose magnitude is driven mainly by the atomic regions where the local Jastrow terms play no role) but rather by its ability to better reproduce local properties in the bonding region. In this respect, Fig. 3 shows that the dependence of the energy curves on the size of the basis set is much well-behaved in the multi-Jastrow case than in the CASSCF one (Fig. 2), particularly at large interatomic distances. Note also that when compared to CASSCF DMC results, the DMC results with the multi-Jastrow nodes are found to display a better convergence as a function of the basis set for all distances.

Another important point is that the value of the dissociation energy (computed as the difference of the total energies at the minimum and at the very large distance, $R=20$ a.u.) is much improved. The value obtained with the 5Z basis set, $D_e = -140.7(4)$ kcal/mol, is very good and is essentially equal to the experimental one within statistical fluctuations.

Finally, regarding the values of r_e and ω_e the values are reasonable but do not seem really improved with the standard FN case. Clearly, to make a more accurate evaluation of these quantities we would need more FN-DMC data around the minimum and larger statistics to decrease the noise; however, such a study is out of the scope of the present work.

TABLE V. Spectroscopic constants of the FH molecule with FN-DMC and various nodal structures. DZ,TZ,QZ, and 5Z refer to the use of the cc-pVnZ basis set with $n=2$, 3, 4, and 5, respectively. CASSCF(2,2) is a CASSCF wavefunction corresponding to two electrons within two orbitals, a two-determinant form here due to symmetry. All distances in atomic units.

Trial wavefunction nodes	R_e (a.u.)	ω_e (cm^{-1})	$E_0(R_{min})$	$E_0(R=20)$	D_0 (kcal/mol)
CASSCF(2,2) DZ basis	1.7042	4508	−100.4245(3)	−100.2099(3)	135.0(3)
CASSCF(2,2) TZ basis	1.7186	4378	−100.4341(3)	−100.2143(4)	138.2(3)
CASSCF(2,2) QZ basis	1.7368	3915	−100.4347(4)	−100.2148(4)	138.3(4)
CASSCF(2,2) 5Z basis	1.7257	3782	−100.4350(4)	−100.2151(4)	138.3(4)
CASSCF(2,2) DZ basis+valence Jastrow	1.7209	3927	−100.4192(3)	−100.2040(3)	135.4(3)
CASSCF(2,2) TZ basis+valence Jastrow	1.7146	4706	−100.4301(4)	−100.2096(4)	138.7(4)
CASSCF(2,2) QZ basis+valence Jastrow	1.7248	4304	−100.4324(4)	−100.2101(5)	139.8(4)
CASSCF(2,2) 5Z basis+valence Jastrow	1.7174	4160	−100.4334(5)	−100.2097(4)	140.7(4)
Exact	1.7326	4137			141.1[a]

[a]Obtained by using the experimental estimate of the correlation energy, $E_c = 0.3882$ a.u. (Ref. 21), the exact atomic energy values (Ref. 20), and the SCF limit for FH (Ref. 18).

VI. COMPUTATIONAL COST OF THE MULTI-JASTROW WAVEFUNCTION

In this section the increase of the computational cost resulting from using the multi-Jastrow wavefunction instead of the standard form is discussed. Let us recall that at each step, the Monte Carlo step (here, VMC or FN-DMC schemes) three basic quantities are to be computed: the trial wavefunction Ψ_T, its first derivatives with respect to each electronic coordinate, $\partial_{i,l}\Psi_T$ ($i=1$ to N_{elec}, $l=x,y,z$), and its Laplacian, $\Sigma_{i,l}\partial_{i,l}^2\Psi_T$. When optimizing the trial wavefunction, additional derivatives are also needed. In the linear method employed here, they include the first derivatives of the trial wavefunction with respect to the parameters to be optimized, denoted as $\partial_{c_p}\Psi_T$, and the Laplacian of these quantities, $\nabla^2(\partial\Psi_T/\partial c_p)$. Note that, in general, the trial wavefunctions are multiconfigurational, that is, are written as a sum of elementary products of up and down determinants. Now, because all the derivatives considered here are linear in terms of these elementary products, we only need to focus our attention on the computation of one given component of the sum. The property of linearity implies that the computational cost of the derivatives will scale as the number of determinants whatever the form of the trial wavefunction considered (either standard or multi-Jastrow). However, this is a very conservative point of view and, in practice, the scaling is much more favorable because computing only once a number of parts common to various determinants allows us to decrease a lot the computational cost. Here, we shall not consider such aspects since they are not related to the fact of using or not a multi-Jastrow form. To summarize, in the standard case, we shall consider, the basic trial wavefunction written as

$$\Psi_T = e^J \Psi_\uparrow \Psi_\downarrow, \qquad (52)$$

where J is a global Jastrow factor such as Eqs. (5) and (6), and Ψ_σ are determinants made of one-particle molecular orbitals. In the multi-Jastrow case we write

$$\Psi_T = \Psi_\uparrow \Psi_\downarrow, \qquad (53)$$

where Ψ_σ ($\sigma=\uparrow$ or $\downarrow$) are determinants made of one-particle molecular orbitals with Jastrow factors attached to them.

Let us now detail the total number of determinants to evaluate the various quantities needed. Next, we shall present actual calculations for the oxygen, the sulfur, and the copper atoms to comment on these computational aspects.

A. The wavefunction, Ψ_T

In the standard case, the computation of the wavefunction involves the evaluation of the common Jastrow, the set of occupied molecular orbitals, $\Psi_{ij}^\sigma = \phi_i(\mathbf{r}_j)$, where $i=1$, n_{orb} (n_{orb} is the total number of molecular orbitals used) and the numerical evaluation of the two determinants, Ψ^σ. In the multi-Jastrow case the only difference lies in the calculation of several individual molecular Jastrows instead of the unique global one. Let us denote the number of different molecular Jastrows used by n_{Jast}. We have $0 \le n_{Jast} \le n_{orb}$ since several molecular orbitals can have the same local Jastrow (in particular, when orbitals are equivalent under

symmetry) and some orbitals can have no Jastrow associated with. When a given Jastrow is attached to a molecular orbital, its value has to be computed for each situation consisting to a different reference electron and for each case a sum over external electrons has to be made. Because of this double sum, the computational cost resulting from the use of a local Jastrow is similar to the cost associated with a global Jastrow. Therefore, the computational cost associated with the Jastrow part is proportional to n_{Jast}.

B. The gradient, $\nabla\Psi_T$

In the standard case, the computation of the gradient of the wavefunction requires the computation of $3N_e$ determinants of matrices obtained by replacing one column of the Slater matrix by a column of derivatives of molecular orbitals. This case can be viewed as special case of our formula, Eq. (26), written as

$$\partial_{i,l}\Psi = \Psi[1 + (P_i^{\sigma_i})_i], \qquad (54)$$

where σ_i is the spin of electron i. The cost of computing $(P_i^{\sigma_i})_i$ being of order N_{σ_i} [see Eq. (24)], the computational cost, T (CPU time), is

$$T \sim 3N_e[N_\uparrow + N_\downarrow]. \qquad (55)$$

In the multi-Jastrow case the evaluation of the gradient is based on the full expression, Eq. (26). In this formula each term $(P_k^\sigma)_k$ requires N_σ operations and, thus, the total computational cost is given by

$$T = 3N_e[N_\uparrow N_\uparrow + N_\downarrow N_\downarrow]. \qquad (56)$$

C. The Laplacian, $\nabla^2\Psi_T$

In the standard case, the computation of the Laplacian of the wavefunction requires the computation of $3N_e$ determinants of matrices obtained by replacing one column of the Slater matrix by a column of second derivatives of molecular orbitals. As in the gradient case, the computational cost is thus given by

$$T \sim 3N_e[N_\uparrow + N_\downarrow]. \qquad (57)$$

In the multi-Jastrow case, we have to distinguish three different cases according to formula (27). A first case corresponds to the third term of Eq. (27) which has already been computed for the gradient and does not cost additional computation. The second and third cases correspond to the cases $k=m$ and $k \ne m$ in Eq. (28). The case $k=m$ has exactly the same structure as previously for the gradient. Accordingly, the cost is given by $T \sim 3N_e[N_\uparrow N_\uparrow + N_\downarrow N_\downarrow]$. In the $k \ne m$ case we need now to consider all pair possibilities. The $(P_k^\sigma)_k$ have already been computed but not the $(P_k^\sigma)_m$. We thus have

$$T \sim 3N_e\left[\frac{N_\uparrow(N_\uparrow - 1)}{2} + \frac{N_\downarrow(N_\downarrow - 1)}{2} \right]. \qquad (58)$$

D. The gradient, $\partial\Psi_T/\partial c_p$

For a given parameter c_p it is necessary to make the derivative of all the lines of the determinant containing this

TABLE VI. Number of determinants to evaluate at each elementary step (for each walker). $N_\uparrow$ and $N_\downarrow$: number of up and down electrons, respectively. $N_e = N_\uparrow + N_\downarrow$: total number of electrons. The quantity $N_{\text{lines}}(\sigma)$, $\sigma = \uparrow, \downarrow$ represents the sum for each parameter to be optimized of the number of lines in the Slater matrix σ containing this parameter, see text.

	Global Jastrow	Multi-Jastrow
	VMC/DMC methods	
Ψ_T	2	2
$\nabla \Psi_T$	$3N_e$	$3N_e^2$
$\nabla^2 \Psi_T$	$3N_e$	$3N_e\left[N_e + \dfrac{N_\uparrow(N_\uparrow - 1)}{2} + \dfrac{N_\downarrow(N_\downarrow - 1)}{2} \right]$
	Optimization step (additional determinants)	
$\dfrac{\partial \Psi_T}{\partial c_p}$	0	$N_{\text{lines}}(\uparrow) + N_{\text{lines}}(\downarrow)$
$\nabla^2 \dfrac{\partial \Psi_T}{\partial c_p}$	0	$3N_e\left[\dfrac{N_\uparrow(N_\uparrow - 1)}{2} + N_\uparrow \right] N_{\text{lines}}(\uparrow) + 3N_e\left[\dfrac{N_\downarrow(N_\downarrow - 1)}{2} + N_\downarrow \right] N_{\text{lines}}(\downarrow)$

parameter. For each of these derivatives, a determinant has to be computed. Let us call N_{opt} the number of determinants to optimize and $N(c_p, \sigma)$ the number of lines containing the parameter c_p in the determinant σ. The total number of determinants, N_{lines}, to calculate is

$$N_{\text{lines}} = N_{\text{lines}}(\uparrow) + N_{\text{lines}}(\downarrow),$$

with

$$N_{\text{lines}}(\sigma) = \sum_{p=1}^{N_{\text{opt}}} N(c_p, \sigma). \tag{59}$$

E. $\nabla^2(\partial \Psi_T / \partial c_p)$

When computing the derivative of Ψ_T with respect to a parameter c_p the derivative of the various lines of the determinants containing this parameter has to be calculated. When making twice the derivative with respect to a given space coordinate, we need to make the derivative with respect to each column. We have to distinguish the case where there is only one column to consider (second derivative), which gives a number of determinants equal to

$$N(c_p, \uparrow)N_\uparrow + N(c_p, \downarrow)N_\downarrow \tag{60}$$

and the case where two different columns are differentiated [there are $([N_\sigma(N_\sigma - 1)]/2)$ such columns], thus giving a number of determinants equal to

$$N(c_p, \uparrow)\frac{N_\uparrow(N_\uparrow - 1)}{2} + N(c_p, \downarrow)\frac{N_\downarrow(N_\downarrow - 1)}{2}. \tag{61}$$

Finally, counting the $3N_e$ components of the Laplacian and the N_{opt} parameters, the total number of determinants to calculate is given by

$$3N_e\left[\frac{N_\uparrow(N_\uparrow - 1)}{2} + N_\uparrow \right] N_{\text{lines}}(\uparrow) + 3N_e\left[\frac{N_\downarrow(N_\downarrow - 1)}{2} + N_\downarrow \right]$$
$$\times N_{\text{lines}}(\downarrow). \tag{62}$$

The various results presented above are summarized in Table VI.

F. Numerical applications and discussion

Now, let us give a quantitative illustration of these various theoretical formulae. For that, we present in Tables VII and VIII some computational times obtained in actual simulations for the three atoms considered here (O, S, and Cu). Timings are given only for the main critical parts of the algorithm, namely, the calculation of the Jastrow part, the matrix elements, the inverse of the Slater matrices, and the determinants. The computational times presented here refer to a series of elementary steps of the algorithm (VMC or optimization step) for a population of 100 walkers. Numbers are given in seconds of CPU time using an ordinary desktop computer. Clearly, only the *relative* values are meaningful here. Note that the present study is intended to illustrate the *gross* features of the computational cost behavior of the algorithm (feasibility, cost as a function of number of electrons, etc.) and not the fine details, such details being too much dependent on the actual implementation of the various formulas and on the specific architecture of the computational platform used.

In Table VII the decomposition of the computational cost for a VMC simulation is presented. Comparisons are made between the standard case where a global Jastrow is used and the new approach where several Jastrow terms are introduced (denoted as $n - J_\phi$). For the oxygen atom, two situations have been considered: a first one where only a single Jastrow has been attached to the innermost $1s$ orbital and a second one where a Jastrow has been attached to each of the five (doubly or singly) occupied orbitals. In the sulfur and copper cases, only three local Jastrow terms have been

considered, corresponding to the choices: $(J_{1s}, J_{2s\text{-}2p}, J_{3s\text{-}3p})$ and $(J_{1s}, J_{2s\text{-}2p\text{-}3s\text{-}3p}, J_{3d\text{-}4s})$ for the sulfur and copper atoms, respectively. Let us begin with the computation of the orbital part of the matrix elements and the inverse of Slater matrices. These two parts are expected to be independent on the use or not of a multi-Jastrow form and this is indeed what is observed. Furthermore, the scaling of the computational cost of these two parts as a function of the number of electrons is found to be in good agreement with theoretical expectations: The various timings reproduce well the fact that the number of matrix elements scales as the square of the number of electrons and the computation of the inverse of small matrices with a law intermediate between a squared and a cubic one (cubic for large enough matrices). Regarding the Jastrow part, the computation of the various properties (zeroth, first, and second derivatives) associated with the global Jastrow term is expected to be of the same order of magnitude than for a typical local Jastrow terms. In the general case the computation should be essentially proportional to the number of such Jastrow terms. These general trends are essentially observed, particularly for the largest system (copper atom) for which the transient effects related to small sizes are minimal. Finally, regarding the calculation of the determinants—the most critical and time-consuming part of such calculations—the numerical results are in agreement with the rapid increase of the number of determinants associated with the multi-Jastrow form. Our theoretical formulas (see Table VI) give an increase by a factor about 14 for the oxygen, 43 for the sulfur, and 127 for the copper atom. These numbers are essentially recovered in our simulations for the biggest systems, namely, the sulfur and copper atoms (factors of about 35 and 105, respectively) with a slight attenuation, thanks to the various tricks presented above for computing the numerous determinants. Note that in the case of the oxygen atom the theoretical factor is not recovered because of the specific finite-size effects related to the very small sizes of the matrices involved.

In Table VIII a similar analysis for the optimization step is presented. The first two columns give the computational costs associated with the optimization of a unique global Jastrow term (standard case). As described in Sec. II A the form employed here for the global Jastrow contains up to 19 parameters. Results are shown for the two extreme cases corresponding to an optimization step involving only one parameter (first column) and the entire set of 19 parameters (second column). The three following columns denoted as $J_\phi[1]\phi$, $J_\phi[2]\phi$, and $J_\phi[12]\phi$ give the various timings corresponding to the optimization of the multi-Jastrow form with a unique Jastrow term J_ϕ attached to a single orbital ϕ for 1, 2 or 12 parameters (maximum number of parameters, see Sec. II B). The three last columns give the timings corresponding to the optimization of two parameters in various situations in order to illustrate a fundamental point discussed below. Regarding the computation of the matrix elements and of the inverse of Slater matrices the situation is exactly the same as in the VMC case previously discussed. The timings for both parts are essentially independent on the nature of the trial wavefunction and on the number of Jastrow terms. Regarding now the Jastrow part, the results presented

have an expected behavior. In both standard and multi-Jastrow cases the computational cost increases with the number of parameters to be optimized and this increase is rather gentle (sublinear) due to the use of common parts which do not need to be recalculated for each parameter. An important point to note is the fact that by using one or two Jastrow terms and only one parameter (column 3, $J_\phi[1]\phi$, or last column, $J_\phi[1]\phi$, $J_\phi[1]\widetilde{\phi}$) we essentially obtain the same timings, a point which illustrates the fact that simultaneous use of a Jastrow term in a multi-Jastrow form does not lead to a sizable increase of the computational cost. Finally, let us end with the computation of determinants, the most time-consuming part of the optimization step. In Table VIII we illustrate the fundamental property of the computational cost of the optimization, namely, the fact that the increase of the number of determinants to evaluate depends only on the quantities $N_{\text{lines}}(\sigma)$, $\sigma = \uparrow, \downarrow$. As already noted, these quantities represent the sum for each parameter of the total number of lines in the Slater matrix σ containing this parameter (see Table VI). Here, this idea is nicely illustrated by showing in the last three columns the various calculations corresponding to the optimization of two parameters with a constant $N_{\text{lines}} = 2$. As observed, the timings are very similar and correspond also to the timings of the column corresponding to $J_\phi[2]\phi$.

VII. SUMMARY AND PERSPECTIVES

In this work a new trial wavefunction suitable for QMC simulations of molecular systems has been presented. This wavefunction is similar to the standard Slater–Jastrow wavefunction except that the common global Jastrow factor is replaced by several Jastrow factors attached to individual molecular orbitals. The physical motivation behind this "multi-Jastrow" form is to allow an orbital-dependent description of the correlation effects in atoms and molecules. Applications to atoms have illustrated that the optimized multi-Jastrow form does not lead to significantly better results for light atoms (first- and second-row elements). However, for a 3d-atom like the copper atom it has been found that by using three different Jastrow terms [for the 1s, (2s-2p-3s-3p), and the (3d-4s) orbitals] a clear improvement of the total energy (with respect to the standard Slater–Jastrow form) can be obtained. Regarding FN results, it has been found that for all the atoms considered the multi-Jastrow wavefunction does not change the FN energies with respect to standard calculations using HF nodes. However, in contrast to atoms, application to the FH molecule has shown that the nodal pattern can be modified and that an almost exact dissociation energy can be obtained. Regarding computational aspects, it has been shown that by using the suitable strategy the computational overhead associated with the use of multi-Jastrow wavefunctions can be kept under reasonable control.

Finally, regarding perspectives, it is clear that this novel form should be implemented and evaluated for much larger molecular systems. In particular, for extended systems it seems natural to put together the use of our Jastrow factors with the use of localized molecular orbitals. By doing that,

TABLE VII. Decomposition of the total computational cost of a series of elementary VMC steps in terms of the main time-consuming parts for the oxygen, sulfur, and copper atoms. Comparison between timings obtained when using a global Jastrow or the multi-Jastrow form with a varying number n of Jastrow factors (denoted as n-J_ϕ). Timings in seconds of CPU time using a standard processor (only the relative values are meaningful). The number of VMC steps has been taken large enough (16 000 steps for 100 walkers) to avoid overhead effects associated with too short simulations.

	Global Jastrow	1-J_ϕ	2-J_ϕ	3-J_ϕ	5-J_ϕ
		Oxygen atom			
Jastrow	29.5	30.2	⋯	⋯	105.2
Matrix elements (orbital part)	18.6	19.2	⋯	⋯	19.8
Inverse of Slater matrices	4.6	4.7	⋯	⋯	4.8
Determinants	1.6	90.5	⋯	⋯	92.4
(No. of determinants)	(50)	(696)	⋯	⋯	(696)
Total	54.3	144.6	⋯	⋯	222.2
		Sulfur atom			
Jastrow	89.5	73.8	128.6	181.1	⋯
Matrix elements (orbital part)	64.7	64.8	64.8	64.7	⋯
Inverse of Slater matrices	18.2	18.0	18.0	17.9	⋯
Determinants	6.2	214.8	214.8	214.7	⋯
(No. of determinants)	(98)	(4226)	(4226)	(4226)	⋯
Total	178.6	371.4	426.2	478.4	⋯
		Copper atom			
Jastrow	329.5	262.2	472.7	677.0	⋯
Matrix elements (orbital part)	335.0	337.5	337.5	337.6	⋯
Inverse of Slater matrices	77.1	75.0	75.1	75.9	⋯
Determinants	20.2	2120.5	2121.2	2123.1	⋯
(No. of determinants)	(174)	(22 098)	(22 098)	(22 098)	⋯
Total	761.8	2795.2	3006.5	3213.6	⋯

TABLE VIII. Decomposition of the total computational cost of a series of elementary optimization steps for the oxygen, sulfur, and copper atoms. Comparison between timings obtained when optimizing the parameters of a global Jastrow and those of a multi-Jastrow form. In this table $J[k]$ corresponds to the optimization of k parameters of a global Jastrow term, $J_\phi[k]\phi$ is the optimization of k parameters of the Jastrow J_ϕ attached to orbital ϕ, $(J_\phi[k]\phi, J_{\tilde\phi}[l]\tilde\phi)$ is the optimization of k parameters of the Jastrow J_ϕ attached to ϕ and l parameters of the Jastrow $\tilde\phi$ attached to $\tilde\phi$. $(J_\phi[k]\phi, J_\phi[k]\tilde\phi)$ is the optimization of k parameters of a single common Jastrow attached to both ϕ and $\tilde\phi$ orbitals. Timings in seconds of CPU time using a standard processor (only the relative values are meaningful). The number of optimization steps has been taken large enough to avoid overhead effects associated with too short simulations.

	$J[1]$	$J[19]$	$J_\phi[1]\phi$	$J_\phi[2]\phi$	$J_\phi[12]\phi$	$J_\phi[2]\phi, J_{\tilde\phi}[0]\tilde\phi$	$J_\phi[1]\phi, J_{\tilde\phi}[1]\tilde\phi$	$J_\phi[1]\phi, J_\phi[1]\tilde\phi$
				Jastrow				
O atom	59.4	92.4	46.0	55.6	124.9	111.4	80.8	46.1
S atom	287.4	382.8	171.7	207.5	547.1	368.4	330.3	175.0
Cu atom	656.3	1291.3	425.2	649.7	3621.6	1277.3	876.9	446.6
				Matrix elements				
O atom	62.0	62.1	62.2	62.1	62.1	62.1	62.1	62.1
S atom	145.9	145.7	145.8	146.5	146.7	146.8	145.7	146.6
Cu atom	372.1	372.5	372.6	371.9	372.2	372.8	372.5	372.5
				Inverse of Slater matrices				
O atom	4.6	4.7	4.7	5.1	4.8	4.7	4.7	4.6
S atom	18.2	17.8	17.9	18.0	18.9	18.4	18.2	18.5
Cu atom	77.8	77.9	78.3	78.2	86.9	80.4	77.7	78.3
				Determinants				
O atom	⋯	⋯	236.9	324.1	767.5	415.1	317.8	307.0
S atom	⋯	⋯	866.9	1187.9	4495.4	1195.6	1138.7	1187.1
Cu atom	⋯	⋯	4629.1	6929.2	32 401.7	6850.7	6761.5	6888.9

we can imagine to define and preoptimize once for all some pairs of local Jastrow and localized molecular orbitals for various chemical situations. By introducing such local elementary pieces, we can hope to be able to build a "molecular Lego" way trial wavefunctions of good-quality for arbitrary complex molecular systems. Work along this direction is presently under way.

ACKNOWLEDGMENTS

[1] R. Assaraf and M. Caffarel, J. Chem. Phys. **119**, 10536 (2003).

[2] D. M. Ceperley and B. J. Alder, J. Chem. Phys. **81**, 5833 (1984).

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

[4] C. A. Coulson and A. H. Neilson, Proc. Phys. Soc. **78**, 831 (1961).

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

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

[7] L. Mitáš, Phys. Rev. Lett. **96**, 240402 (2006).

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

[9] F. A. Matsen, Adv. Quantum Chem. **1**, 59 (1964).

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

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

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

[13] B. L. Hammond, W. A. Lester Jr., and P. J. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry*, World Scientific Lecture and Course Notes in Chemistry Vol.1 (World Scientific, Singapore, 1994).

[14] See, e.g., W. H. Press, B. P. Flannery, S. A. Teukolsky, and W. T. Vetterling, *Numerical Recipes in C: The Art of Scientific Computing* (Cambridge University Press, 1992).

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

[16] C. F. Bunge, J. A. Barrientos, A. Viver Bunge, and J. A. Cogordan, Phys. Rev. A **46**, 3691 (1992).

[17] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).

[18] H. Müller, R. Franke, S. Vogtner, R. Jaquet, and W. Kutzelnigg, Theor. Chem. Acc. **100**, 85 (1998).

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

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

[21] S. J. Chakravorty and E. R. Davidson, J. Phys. Chem. **100**, 6167 (1996); R. S. Grev, C. L. Janssen, and H. F. Schaefer, J. Chem. Phys. **95**, 5128 (1991).