Beyond the locality approximation in the standard diffusion Monte Carlo method

Michele Casula

Department of Physics, University of Illinois at Urbana-Champaign, 1110 West Green St, Urbana, Illinois 61801, USA

We present a way to include nonlocal potentials in the standard diffusion Monte Carlo method without using the locality approximation. We define a stochastic projection based on a fixed node effective Hamiltonian, whose lowest energy is an upper bound of the true ground-state energy, even in the presence of nonlocal operators in the Hamiltonian. The variational property of the resulting algorithm provides a stable diffusion process, even in the case of divergent nonlocal potentials, like the hard-core pseudopotentials. It turns out that the modification required to improve the standard diffusion Monte Carlo algorithm is simple.

The diffusion Monte Carlo (DMC) algorithm is one of the most successful methods to compute the ground-state properties of quantum systems. Although the fixed node (FN) approximation is needed to cure the infamous sign problem for fermions, the accuracy of the DMC framework has yielded many benchmark results.[1] However, when the DMC method is applied to *ab initio* realistic Hamiltonians, its computational cost scales proportionally to $Z^{6.5}$, where Z is the atomic number.[2] Therefore, the use of pseudopotentials is necessary to make those calculations feasible.

Since the pseudopotentials are usually nonlocal, the "locality approximation" is made besides the FN approximation, by replacing the true Hamiltonian H with an *effective* one H^{eff}, which reads[3]

$$H^{\text{eff}} = K + V_{\text{loc}} + \frac{\int dx' \langle x' | V_{\text{nonloc}} | x \rangle \Psi_T(x')}{\Psi_T(x)}, \quad (1)$$

where K is the kinetic operator, V_{loc} is the local potential, and the last term in Eq. (1) is the nonlocal potential localized by means of the trial wave function Ψ_T. The projection is then realized by iteratively applying the operator $G = \exp[-\tau(H^{\text{eff}} - E_{\text{eff}})]$ to Ψ_T in order to filter out its high-energy components. The localized potential enters in the branching part (birth and death processes) of the algorithm, while the usual FN constraint is employed to limit the diffusion process within the nodal pockets of Ψ_T, and avoid the fermionic sign problem. Thus E_{eff} is the FN ground-state energy of H^{eff}, computed during the sampling of the mixed distribution $\Psi_{\text{eff}} \Psi_T$:

$$E_{\text{eff}} = \frac{\langle \Psi_{\text{eff}} | H^{\text{eff}} | \Psi_T \rangle}{\langle \Psi_{\text{eff}} | \Psi_T \rangle} = \frac{\langle \Psi_{\text{eff}} | H | \Psi_T \rangle}{\langle \Psi_{\text{eff}} | \Psi_T \rangle} = E_{\text{MA}}. \quad (2)$$

E_{MA} is the *mixed average* of H, and the above identity holds because $H^{\text{eff}} \Psi_T / \Psi_T = H \Psi_T / \Psi_T$. Since Ψ_{eff} is the FN ground-state of H^{eff}, which differs from H, E_{MA} is no longer equal to the variational FN energy of H, defined as

$$E_{\text{FN}} = \langle \Psi_{\text{eff}} | H | \Psi_{\text{eff}} \rangle / \langle \Psi_{\text{eff}} | \Psi_{\text{eff}} \rangle. \quad (3)$$

Therefore, in contrast with the case of local Hamiltonians, E_{MA} calculated with the locality approximation does not in

general give an upper bound to the ground-state energy of H (variational principle).

In previous work,[4] we introduced the lattice regularized diffusion Monte Carlo (LRDMC) algorithm, which provides an upper bound for the true ground-state energy and allows estimation of E_{FN}, even in the case of nonlocal potentials. In this paper we propose an extension of the standard DMC framework that gives the same results as the LRDMC method, after a proper modification of the DMC propagator.

We start by considering the importance sampling Green's function

$$G(x' \leftarrow x, \tau) = \frac{\Psi_T(x')}{\Psi_T(x)} \langle x' | e^{-\tau(H - E_T)} | x \rangle, \quad (4)$$

where E_T is an energy offset, τ the time step, and x a vector of particle coordinates. In the diffusion Monte Carlo method, $G(x' \leftarrow x, \tau)$ is iteratively applied to Ψ_T^2, in order to sample stochastically the mixed distribution $\Phi(x, t) = \Psi_T(x) \Psi(x, t)$, $\Psi(x, t)$ converging to the lowest possible state in energy. To rewrite $G(x' \leftarrow x, \tau)$ [Eq. (4)] in a practical way, it is necessary to resort to the Trotter breakup, which is exact in the limit of $\tau \rightarrow 0$. Here we split the Hamiltonian into local and nonlocal operators, and we end up with the following expression for the Green's function:

$$G(x' \leftarrow x, \tau) \simeq \int dx'' T_{x', x''}(\tau) G_{\text{DMC}}(x'' \leftarrow x, \tau), \quad (5)$$

where $G_{\text{DMC}}(x' \leftarrow x, \tau)$ is the usual DMC propagator,[1]

$$\frac{1}{(2\pi\tau)^{3N/2}} \exp\left(-\frac{[x' - x - \tau v(x)]^2}{2\tau}\right) e^{-\tau[E_L^{\text{loc}}(x') - E_T]}, \quad (6)$$

and $T_{x', x}(\tau)$ is the matrix containing the nonlocal potential,

$$\frac{\Psi_T(x')}{\Psi_T(x)} \langle x' | e^{-\tau V_{\text{nonloc}}} | x \rangle \simeq \delta_{x', x} - \tau V_{x', x}. \quad (7)$$

In the above equations N is the total number of particles, $v(x) = \nabla \ln|\Psi_T(x)|$ the drift velocity, $E_L^{\text{loc}}(x) = (K + V_{\text{loc}}) \Psi_T(x) / \Psi_T(x)$ the contribution to the local energy coming from the local operators, and $V_{x', x} = \frac{\Psi_T(x')}{\Psi_T(x)} \langle x' | V_{\text{nonloc}} | x \rangle$. The final form of G_{DMC} has been obtained by further splitting the Hamiltonian into kinetic and potential parts, while

the exponential of the nonlocal potential in T has been linearized up to order τ.

If the case of pseudopotentials, the number of nonzero matrix elements $V_{x',x}$ will be *finite*, once a quadrature rule with a discrete mesh of points is applied to evaluate the projection over the angular components of the pseudopotential.[3,5] Therefore, the process in $G(x' \leftarrow x, \tau)$ driven by $T_{x',x}(\tau)$ can be calculated using a heat bath algorithm, since $T_{x',x}(\tau)/\Sigma_{x''}T_{x'',x}(\tau)$ can be seen as a transition probability, and it can be computed *a priori* for all possible new coordinates x'. We notice that the matrix elements $T_{x',x}(\tau)$ are easily evaluated in a standard DMC algorithm, since $V_{x',x}$ are already computed to calculate the localized pseudopotential in Eq. (1):

$$\frac{\int dx' \langle x' | V_{\text{nonloc}} | x \rangle \Psi_T(x')}{\Psi_T(x)} = \sum_{x'} V_{x',x}. \qquad (8)$$

At variance with the locality approximation, $V_{x',x}$ contribute now to move the particles, according to the transition matrix T (T moves).

An important limitation of this idea is given by the sign problem. Indeed both $\Psi_T(x')/\Psi_T(x)$ and $\langle x' | V_{\text{nonloc}} | x \rangle$ can change sign, which should be included in the weights, but this yields averages with exponentially increasing noise. A solution is to apply the FN approximation not only to G_{DMC} but also to T, which becomes

$$T_{x',x}^{\text{FN}}(\tau) = \delta_{x,x'} - \tau V_{x',x}^-, \qquad (9)$$

where we defined $V_{x',x}^{\pm} = (1/2)(V_{x',x} \pm |V_{x',x}|)$. In practice, we keep only those matrix elements which give a positive $T_{x',x}(\tau)$. Moreover, we add to the diagonal potential the so-called sign flip term, i.e., the sum over the discarded matrix elements $V_{x',x}^+$. Therefore, the local potential becomes

$$V^{\text{eff}}(x) = V_{\text{loc}}(x) + \sum_{x'} V_{x',x}^+. \qquad (10)$$

This is equivalent to working with a new effective FN Hamiltonian

$$H_{x,x}^{\text{eff}} = K + V^{\text{eff}}(x),$$

$$H_{x',x}^{\text{eff}} = \langle x' | V_{\text{nonloc}} | x \rangle \quad \text{if } V_{x',x} < 0. \qquad (11)$$

In contrast to the effective Hamiltonian of the locality approximation written in Eq. (1), the ground-state energy E_{eff} ($=E_{\text{MA}}$) of the above H^{eff} is an upper bound for the ground-state energy of the true H. As shown in Ref. 6 for the lattice Green's function Monte Carlo method, this variational property is due to the sign flip term (positive contribution) added to the local potential, and the T moves driven by the off diagonal matrix elements $V_{x',x}^-$. Instead, in the locality approximation also $V_{x',x}^-$ is summed in the diagonal part [Eq. (8)], and this leads to an attractive potential, which cannot provide a variational property for E_{MA}. Moreover, we found that the negative divergences of the fully localized potential on the nodes of Ψ_T are responsible in some case (e.g., see
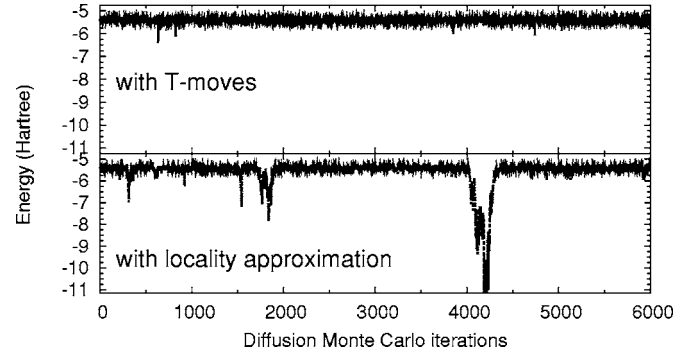

FIG. 1. Energies for the carbon pseudoatom with $\tau = 0.08H^{-1}$ at the given DMC generation. Ψ_T is an antisymmetrized geminal power wave function with a three-body Jastrow factor (Refs. 12 and 15). We report results for the locality approximation ($H^{\alpha,\gamma}$ with $\alpha = 1$ and $\gamma = 0$) and the algorithm with T moves ($\alpha = 0$, $\gamma = 0$).

Fig. 1) for numerical instabilities in the locality approximation, which disappear once H^{eff} in Eq. (11) is used together with the T^{FN} moves. Indeed, whenever $V_{x',x}^-$ is large, it pushes the walker away from the attractive regions of the localized potential, and protects the sampling from divergences in the weights.

Once a T^{FN} move is generated according to the transition probability $T_{x',x}^{\text{FN}}(\tau)/\Sigma_{x''}T_{x'',x}^{\text{FN}}(\tau)$, the walker should acquire the weight $w_T(x, \tau) = \Sigma_{x''}T_{x'',x}^{\text{FN}}(\tau)$ due to the normalization of the T^{FN} matrix. This weight can be recast as an exponential form valid up to order τ,

$$w_T = 1 - \tau \sum_{x'} V_{x',x}^- \simeq \exp\left(-\tau \sum_{x'} V_{x',x}^-\right). \qquad (12)$$

Thus the overall weight $w(x, \tau)$ of $G(x' \leftarrow x, \tau)$ will be

$$w(x, \tau) = w_{\text{DMC}}w_T = \exp[-\tau(E_L(x) - E_T)], \qquad (13)$$

where w_{DMC} is the weight of G_{DMC} for the effective Hamiltonian (V_{loc} replaced by V^{eff}), and $E_L(x) = H^{\text{eff}}\Psi_T/\Psi_T = H\Psi_T/\Psi_T$ is the local energy. Notice that a nonsymmetric branching factor has been included in G_{DMC} [Eq. (6)]. When we use the exponential form in Eq. (12), and consequently the weight in Eq. (13), the time step error is usually smaller than that obtained with the linear form. This can be understood in the limit of perfect importance sampling. Indeed, if Ψ_T is close to the ground state of H, the weight in Eq. (13) is almost constant, since the variance of $E_L(x)$ is small, and the time step bias is reduced.

The proposed DMC scheme for fixed node Hamiltonians with nonlocal potentials is the following: (i) perform a diffusion-drift move according to $G_{\text{diff}}(x' \leftarrow x, \tau) = \exp[-(x' - x - \tau v(x))^2/2\tau]/(2\pi\tau)^{3N/2}$ as is done in the standard DMC algorithm, and accept or reject this move according to the probability

$$\min\left(1, \frac{G_{\text{diff}}(x \leftarrow x', \tau)\Psi_T^2(x')}{G_{\text{diff}}(x' \leftarrow x, \tau)\Psi_T^2(x)}\right); \qquad (14)$$

(ii) weight the walker with the factor $\exp[-\tau(E_L(x') - E_T)]$; (iii) displace the walker a second time, with a T move se-

lected according to the transition probability $p(x'' \leftarrow x', \tau) = T^{FN}_{x'',x'}(\tau)/\Sigma_y T^{FN}_{y,x'}(\tau)$, computed *a priori* for all possible new x''. The branching process will be the same as in the usual DMC algorithm. In practice, only the T move is the new step, which is performed *after* weighting the walker.[7] Although we perform an acceptance or rejection step [Eq. (14)], which has been shown to reduce the time step error[8] in G_{DMC}, the algorithm does not satisfy exactly the detailed balance except in the limit of $\tau \to 0$, due to the break up of G into G_{DMC} and T^{FN} [Eq. (5)], and the use of a nonsymmetric branching factor in Eq. (6).

In order to estimate the variational FN energy E_{FN} [Eq. (3)], and study the quality of the locality approximation, we introduce a more general effective Hamiltonian[4] $H^{\alpha,\gamma}$,

$$H^{\alpha,\gamma}_{x,x} = K + V_{loc}(x) + (1+\gamma)\sum_{x'} V^+_{x',x} + \alpha(1+\gamma)\sum_{x'} V^-_{x',x},$$

$$H^{\alpha,\gamma}_{x',x} = -\gamma\langle x'|V_{non\ loc}|x\rangle \quad \text{if } V_{x',x} > 0,$$

$$H^{\alpha,\gamma}_{x',x} = [1 - \alpha(1+\gamma)]\langle x'|V_{non\ loc}|x\rangle \quad \text{if } V_{x',x} < 0, \quad (15)$$

where $0 \leq \alpha \leq 1$ and $0 \leq \gamma \leq 1/\alpha - 1$ are two external parameters. In order to sample the Green's function $G(x' \leftarrow x, \tau)$ for $H^{\alpha,\gamma}$, it is sufficient to modify the matrix $T_{x',x}(\tau)$, which becomes

$$T^{\alpha,\gamma}_{x',x} = \begin{cases} 1 & \text{if } x = x', \\ \tau\gamma V^+_{x',x} & \text{if } V_{x',x} > 0, \\ -\tau[1 - \alpha(1+\gamma)]V^-_{x',x} & \text{if } V_{x',x} < 0. \end{cases} \quad (16)$$

The ground state $E(\alpha,\gamma)$ of $H^{\alpha,\gamma}$ is equal to $E_{MA}(\alpha,\gamma)$ [Eq. (2)], since $H^{\alpha,\gamma}\Psi_T/\Psi_T = H\Psi_T/\Psi_T$ by construction. The Hamiltonian in Eq. (11) is recovered with $\alpha=0$ and $\gamma=0$, while the Hamiltonian of the locality approximation [Eq. (1)] is obtained with $\alpha=1$ and $\gamma=0$. Therefore, $H^{\alpha,\gamma}$ can interpolate between these two extremes, but the variational principle for $E_{MA}(\alpha,\gamma)$ is not guaranteed as soon as $\alpha \neq 0$, since the attractive term $\alpha(1+\gamma)\Sigma_{x'}V^-_{x',x}$ is added to the diagonal potential. However by means of $H^{\alpha,\gamma}$ one can estimate the value of $E_{FN}(\alpha,\gamma)$ [Eq. (3)], which is *variational* for every α and γ, since it is the expectation value of the true H on the ground state of $H^{\alpha,\gamma}$. Indeed $H = H^{\alpha,\gamma} - (1+\gamma)\partial_\gamma H^{\alpha,\gamma}$, and the Hellmann-Feynman theorem leads to the relation

$$E_{FN}(\alpha,\gamma) = E(\alpha,\gamma) - (1+\gamma)\partial_\gamma E(\alpha,\gamma). \quad (17)$$

One can show[9] that, for a given value of α, the lowest $E_{FN}(\alpha,\gamma)$ is obtained for $\gamma=0$. Therefore, in order to find the best variational estimate of the ground state of H, it is enough to calculate the expression in Eq. (17) with $\gamma=0$. In this way one can check which α provides the best variational state for H. The derivative $\partial_\gamma E(\alpha,0)$ can be computed with either finite differences or correlated sampling. In both cases, one should keep in mind that $\gamma < 1/\alpha - 1$, to guarantee the positivity of the $T^{\alpha,\gamma}$ matrix [Eq. (16)], and so calculating $E_{FN}(\alpha,0)$ becomes harder as α gets closer to 1.

Here we present the application of the method to the Si and C pseudoatoms. We computed $E_{MA}(\alpha,0)$ and $E_{FN}(\alpha,0)$
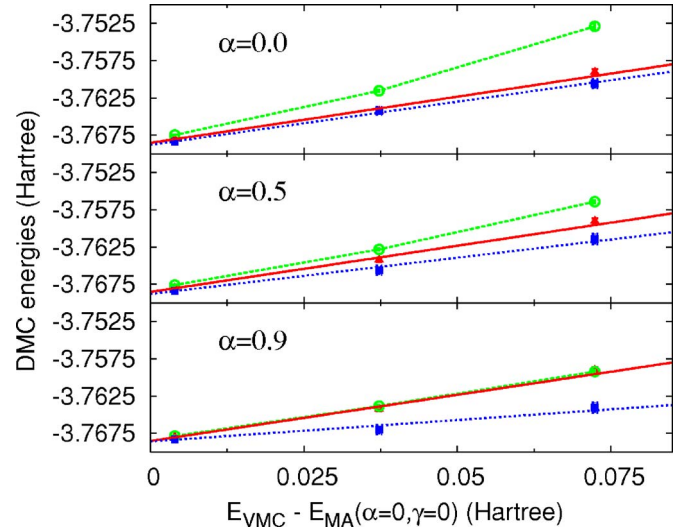

FIG. 2. (Color online) $E_{MA}(\alpha,0)$ (green, dashed line) and $E_{FN}(\alpha,0)$ (blue, dotted line) energies for the silicon pseudoatom with different values of the effective Hamiltonian parameter α. The DMC energies with locality approximation (red, solid line), corresponding to $E_{MA}(1,0)$, are reported in all panels for reference. We used three different Ψ_T's, which have the same determinantal part. A more accurate Ψ_T corresponds to a smaller difference between its variational energy (E_{VMC}) and the $E_{MA}(0,0)$ energy, reported on the abscissa.

for $\alpha=0$, 0.5, 0.9, and the DMC energy with the locality approximation, which corresponds to $E_{MA}(1,0)$. With the aim of quantifying the locality error, and the correction provided by the effective Hamiltonian $H^{\alpha,\gamma}$, we used three different trial wave functions [with no Jastrow, a two-body, and a three-body (electron-electron-ion) Jastrow factor, respectively], sharing the same determinantal part, and hence the same nodes. In this way, the FN error can be separated from the effect of the locality approximation, which causes a dependence of the DMC energy on the shape of Ψ_T.

For the Si atom we used an *s-p* norm-conserving Hartree-Fock (HF) pseudopotential, which is soft and has been generated using the Vanderbilt construction.[10] The determinantal part of Ψ_T is a HF wave function with a $6s6p/1s1p$ Gaussian basis set. The two-body Jastrow factor is from Ref. 11, while the three-body Jastrow factor is from Ref. 12. Both of the Jastrow factors have been optimized using an energy minimization procedure.[13] The results are reported in Fig. 2. The variational $E_{FN}(\alpha,0)$ improves going from $\alpha=0$ to $\alpha=0.9$, i.e., approaching the locality approximation. It means that at least for this soft pseudopotential the locality approximation ($\alpha=1$, $\gamma=0$ gives a ground state which is a good variational wave function for H. Notice, however, that the standard DMC energies $E_{MA}(1,0)$ have a sizable locality error, while E_{FN} with $\alpha=0.9$ depends only slightly on the shape of the trial wave function. A similar result was obtained with the LRDMC method for the same pseudoatom.[4]

For the C atom we chose to work with a Stevens-Basch-Krauss pseudopotential,[14] which is extremely hard, since it diverges like $1/r^2$ in the *s* channel, and $1/r$ in its local component. The Slater part of Ψ_T is an antisymmetrized geminal power wave function[12,15] with a $2s2p$ Gaussian basis set,
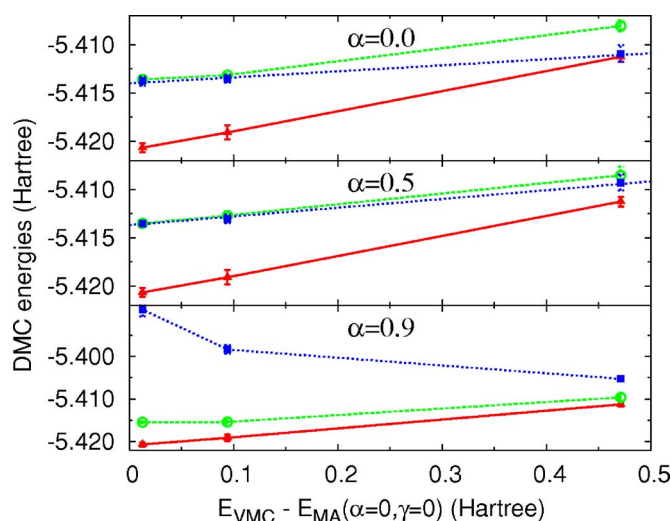
FIG. 3. (Color online) The same as Fig. 2, but for the carbon pseudoatom.

optimized in the presence of the three-body Jastrow factor by minimizing its variational energy.[13] The determinantal part has been kept fixed in the other two Ψ_T's, which differ only by their Jastrow factors. The results are plotted in Fig. 3. Here the locality approximation is very poor, as it leads to nonvariational E_{MA}. The spikes in Fig. 1, coming from regions of the configuration space where the effective potential is attractive, are surely responsible for the nonvariational results. Surprisingly, Ψ_T without the Jastrow factor, which has a higher energy, leads to much more stable DMC simula-

tions. The locality approximation, which relies on the quality of the shape of Ψ_T in the core, performs poorly with this hard-core pseudopotential, since it is difficult to find the optimal shape of Ψ_T in the core region, due to the divergence of the nonlocal pseudopotential. Indeed $E_{FN}(\alpha,0)$ is higher for $\alpha=0.9$, being worst for the three-body Jastrow factor. On the other hand, the best variational $E_{FN}(\alpha,0)$ is obtained for $\alpha=0$, irrespective of the form of the Jastrow factor.

To summarize, we have described a scheme to treat nonlocal potentials within the standard DMC method. We have extended the DMC formalism to handle a generic Hamiltonian with discrete off-diagonal matrix elements and the fixed node approximation. Only a simple modification of the standard algorithm is required to include the T moves generated according to the nonlocal potentials. By using an effective Hamiltonian approach, we showed that it is possible to have stable simulations, even in the case of divergent hardcore pseudopotentials, and obtain variational results. A similar effective Hamiltonian has been successfully used in the LRDMC method. The difference is in the kinetic part, which is discretized in the lattice regularized approach. The LRDMC and the DMC methods have the same efficiency for small Z_{eff}, although it is possible to have a gain in the LRDMC efficiency by an *ad hoc* choice of the kinetic parameters, particularly for heavier elements. We conclude, by noting that the same Green's function presented here can be used in the reptation quantum Monte Carlo method.[16]

[1] W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001).
[2] D. M. Ceperley, J. Stat. Phys. **43**, 815 (1986); A. Ma, N. D. Drummond, M. D. Towler, and R. J. Needs, Phys. Rev. E **71**, 066704 (2005).
[3] L. Mitáš, E. L. Shirley, and D. M. Ceperley, J. Chem. Phys. **95**, 3467 (1991).
[4] M. Casula, C. Filippi, and S. Sorella, Phys. Rev. Lett. **95**, 100201 (2005).
[5] S. Fahy, X. W. Wang, and Steven G. Louie, Phys. Rev. B **42**, 3503 (1990).
[6] D. F. B. ten Haaf, H. J. M. van Bemmel, J. M. J. van Leeuwen, W. van Saarloos, and D. M. Ceperley, Phys. Rev. B **51**, 13039 (1995).
[7] Since the normalization of $T_{x',x}^{FN}$ [Eq. (12)] depends on x, the best choice is to weight the walker before the T move.
[8] P. J. Reynolds *et al.*, J. Chem. Phys. **77**, 5593 (1982); C. J. Umrigar *et al.*, *ibid.* **99**, 2865 (1993).
[9] S. Sorella, cond-mat/0201388 (unpublished).
[10] D. Vanderbilt, Phys. Rev. B **32**, 8412 (1985).
[11] C. Filippi and C. J. Umrigar, J. Chem. Phys. **105**, 213 (1996).
[12] M. Casula, C. Attaccalite, and S. Sorella, J. Chem. Phys. **121**, 7110 (2004).
[13] S. Sorella, Phys. Rev. B **71**, 241103(R) (2005).
[14] W. J. Stevens, H. Basch, and M. Krauss, J. Chem. Phys. **81**, 6026 (1984).
[15] M. Casula and S. Sorella, J. Chem. Phys. **119**, 6500 (2003).
[16] S. Baroni and S. Moroni, Phys. Rev. Lett. **82**, 4745 (1999).