JCTC Journal of Chemical Theory and Computation

Article

pubs.acs.org/JCTC

Practical Schemes for Accurate Forces in Quantum Monte Carlo

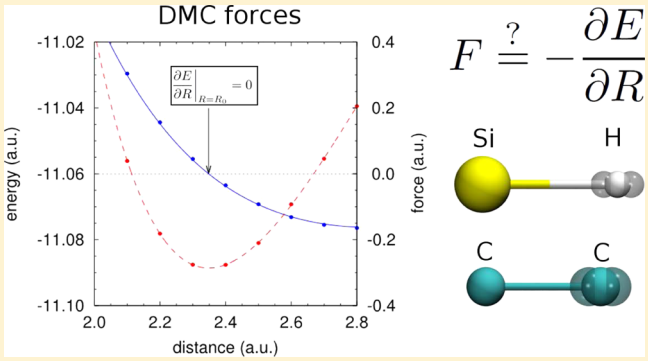
S. Moroni,*[,†] S. Saccani,[‡] and C. Filippi*[,§]

[†]CNR-IOM DEMOCRITOS, Istituto Officina dei Materiali, and SISSA Scuola Internazionale Superiore di Studi Avanzati, Via Bonomea 265, I-34136 Trieste, Italy

[‡]SISSA Scuola Internazionale Superiore di Studi Avanzati, Via Bonomea 265, I-34136 Trieste, Italy

[§]MESA+ Institute for Nanotechnology, University of Twente, P.O. Box 217, 7500 AE Enschede, The Netherlands

ABSTRACT: While the computation of interatomic forces has become a well-established practice within variational Monte Carlo (VMC), the use of the more accurate Fixed-Node Diffusion Monte Carlo (DMC) method is still largely limited to the computation of total energies on structures obtained at a lower level of theory. Algorithms to compute exact DMC forces have been proposed in the past, and one such scheme is also put forward in this work, but remain rather impractical due to their high computational cost. As a practical route to DMC forces, we therefore revisit here an approximate method, originally developed in the context of correlated sampling and named here the Variational Drift-Diffusion (VD) approach. We thoroughly investigate its accuracy by checking the consistency between the approximate VD force and the derivative of the DMC potential energy surface for the SiH and C_2 molecules and employ a wide range of wave functions optimized in VMC to assess its robustness against the choice of trial function. We find that, for all but the poorest wave function, the discrepancy between force and energy is very small over all interatomic distances, affecting the equilibrium bond length obtained with the VD forces by less than 0.004 au. Furthermore, when the VMC forces are approximate due to the use of a partially optimized wave function, the DMC forces have smaller errors and always lead to an equilibrium distance in better agreement with the experimental value. We also show that the cost of computing the VD forces is only slightly larger than the cost of calculating the DMC energy. Therefore, the VD approximation represents a robust and efficient approach to compute accurate DMC forces, superior to the VMC counterparts.



1. INTRODUCTION

Quantum Monte Carlo (QMC) methods are a set of numerical tools to calculate ground state properties, selected excitations, and thermal averages of many-body systems and models with good accuracy and a computational cost which, albeit typically large, scales favorably with the system size. For the electronic structure problem, the most widely used QMC methods are Variational Monte Carlo (VMC) and Fixed-Node Diffusion Monte Carlo (FN-DMC).[1,2] In VMC, one calculates expectation values on a trial function Ψ—an explicitly known correlated wave function deemed a good approximation to the quantum state of interest—using a Metropolis algorithm. In FN-DMC, one solves the Schrödinger equation subject to the constraint that the nodes of the solution coincide with the nodes of Ψ, using a stochastic implementation of the power method. For given Ψ, DMC is less efficient, but more accurate, than VMC.

A major issue that hinders[2,3] a more widespread application of QMC methods to the electronic structure problem is the alleged lack of efficient and reliable calculation of the forces acting on the nuclei. We consider "efficient" a force calculation whose computational cost is within a sufficiently small factor of the calculation of the energy, this factor being independent of the system size.[4] We consider "reliable" a force calculation whose result is in sufficiently good agreement with the slope of the potential energy surface (PES) calculated from the energy. Furthermore, a practical force calculation should be automatic (e.g., no human check, postprocessing or input at intermediate steps in a loop over nuclear geometries).

It seems fair to say that these conditions are by now met in VMC. Modern energy minimization techniques[5−7] offer a robust and stable way for optimizing all the variational parameters in the trial function. Exploiting the ensuing simplification in the VMC force calculation, structural optimizations in the ground and excited states,[8,9] chemical reaction paths,[10,11] and even molecular dynamics of high-pressure liquid hydrogen[12] have been reported. We note that highly parametrized, fully optimized trial functions often give better results at VMC level than simple trial functions at DMC level.

The situation is less satisfactory in DMC. A scheme that guarantees full consistency between force calculation and PES has been recently proposed,[13] and we introduce another such method in section 2.2. Unfortunately, neither can be considered efficient. For instance, our scheme features a strong square-root

Received: June 26, 2014
Published: September 22, 2014

 © 2014 American Chemical Society dx.doi.org/10.1021/ct500780r | J. Chem. Theory Comput. 2014, 10, 4823−4829

dependence on the time step, very difficult to extrapolate out, as shown in Figure 1. We also mention that structural

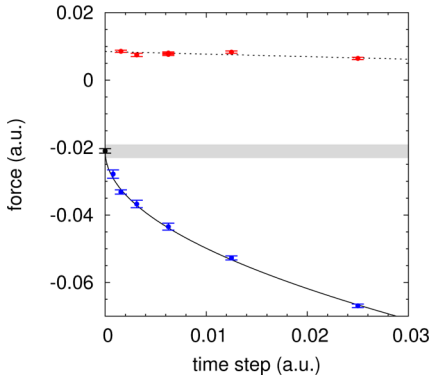


Figure 1. Force for the carbon dimer with the atoms 2.4 au apart as a function of the time step, calculated in DMC with an extremely poor wave function (HF_{1pJ2}, see section 3 and Table 1). Red points and dotted line: approximate algorithm.[15] Blue points and solid line: exact algorithm; the extrapolation to zero time step is shown by the black point. The shaded area shows the range within one standard deviation of the force obtained from the PES.

optimization in DMC without force calculations has been demonstrated,[14] but systematic applications of this approach, to our knowledge, have yet to appear.

Relaxing the requirement of strict consistency between forces and PES, several different schemes have been proposed. The approximation $\nabla\Phi/\Phi \simeq \nabla\Psi/\Psi$[16] leads to an efficient calculation,[17,18] but the error incurred can be significant.[19] Low-variance Hellmann−Feynman forces have been developed,[20−22] but they neglect a nodal contribution, which again can be significant.[19] Such nodal terms are approximately evaluated in ref 19, using the future-walking method,[23] but future-walking quickly becomes impractical as the system size increases.

In this work, we reconsider in detail an approximate method introduced in ref 15 in the context of a correlated sampling scheme for DMC. Based on tests for the SiH and the C_2 dimer, the method indeed qualifies as efficient and (with all but the poorest trial functions) reliable. Figure 2 compares the force and the slope of the PES for the C_2 dimer. The discrepancy is negligible in determining equilibrium bond length and hardly detectable elsewhere, and the computational cost of the force calculation is not much higher than that of the energy. Although the results reported here have been obtained using finite increments, the differential implementation of this approximation is no more difficult than for VMC. Therefore, we expect that a good scaling of the computational cost with the system size can be achieved.[24] This method has been successfully used in structural optimizations[8] and chemical reaction paths.[10]

In the process, we also present instructive results obtained in VMC with various trial functions at different levels of optimization.

2. CALCULATION OF FORCES

In QMC the energy is calculated as

$$E = \int dR E_L(R) P(R) \equiv \langle E_L \rangle \tag{1}$$

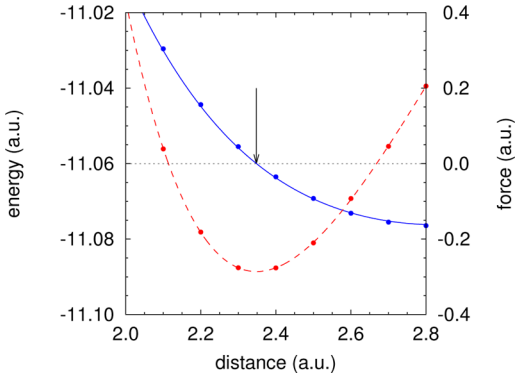


Figure 2. Energy (left scale, dashed line, red online) and force (right scale, solid line, blue online) of the carbon dimer as a function of the atomic distance. The points are FN-DMC results with a good wave function (CAS_{all}, see text). The forces are calculated with the algorithm of ref 15. The statistical errors are much smaller than the symbol size. The dashed curve, a polynomial fit to the calculated energies, represents the potential energy surface (PES), and the solid curve is minus the gradient of the PES. The arrow shows the experimental value of the equilibrium bond length. The discrepancy between the forces obtained[25] from the PES and those directly calculated is statistically significant but barely visible on this scale.

where R is the coordinates of the N electrons, $E_L(R) = \langle R|H|\Psi\rangle/\langle R|\Psi\rangle$ is the local energy, and $P(R)$ is the probability distribution sampled in the Monte Carlo run. In VMC $P(R) = \Psi^2(R)/\langle\Psi|\Psi\rangle$, and in DMC $P(R) = \Psi(R)\Phi(R)/\langle\Psi|\Phi\rangle$, where $\Phi(R)$ is the FN solution. The force is

$$F = -\nabla_\alpha E = -\langle \nabla_\alpha E_L + (E_L - E)\nabla_\alpha \ln P \rangle \tag{2}$$

where the gradient is taken with respect to the nuclear coordinates R_α.

2.1. Variational Monte Carlo. In VMC, both E_L and P are known functions of R and of the nuclear coordinates, and differentiation can be carried out explicitly.[24,26] In particular, using adjoint algorithmic differentiation,[24] the computational cost of the force for a given electron configuration is larger than that of the energy by a modest factor independent of N.

The estimator (eq 2) has manifestly zero variance if both the trial function and its gradient are exact. With a necessarily approximated trial function, E_L diverges at the nodes of Ψ, and so does the variance of F.[12] This occasionally produces large spikes in the force, detrimental to iterative procedures such as structural optimization. The divergence of the variance of F is removed by sampling the square of a modified trial function $\tilde{\Psi}(R)$, which differs from Ψ only near the nodes, where Ψ vanishes but $\tilde{\Psi}$ stays finite—and by weighting the averages with the factor $\Psi^2(R)/\tilde{\Psi}^2(R)$.[12] The variance of F is further reduced using the space-warp transformation of refs 15 and 27. We note that, in the case of an all-electron calculation, the divergence of the electron−nucleus potential also results in a divergence of the variance of F, which can be cured as described in refs 21 and 22 or using the space-warp transformation.

The energy E depends on the nuclear coordinates not only explicitly but also implicitly through the variational parameters. The force F therefore consists of two terms

$$\frac{dE}{dR_\alpha} = \frac{\partial E}{\partial R_\alpha} + \sum_i \frac{\partial E}{\partial c_i}\frac{dc_i}{dR_\alpha} \tag{3}$$

where c_i are the parameters in Ψ, as detailed in ref 19. The second term vanishes when the trial function is fully optimized

dx.doi.org/10.1021/ct500780r | J. Chem. Theory Comput. 2014, 10, 4823−4829

by energy minimization,[8,12] but it is often assumed to be negligibly small even when the optimization is incomplete.[17,19] A typical instance of this approximation is the choice of mean-field orbitals in the Slater determinant times an optimized Jastrow factor. We will show explicitly the effect of neglecting the orbital optimization in section 4.

2.2. Diffusion Monte Carlo. In DMC, there is the additional difficulty that the unnormalized distribution $\tilde{P} = \Phi\Psi$ is not known in closed form but is sampled through a stochastic implementation of the power method, whose n-th iteration is defined by the integral equation

$$\tilde{P}^{(n)}(R') = \int \mathrm{d}R G(R', R)\tilde{P}^{(n-1)}(R) \tag{4}$$

Here, $G(R',R) = \Psi(R')\langle R'|\exp[-(H - E_T)\tau]|R\rangle/\Psi(R)$ is the importance-sampled Green function, τ is the time step, and E_T is an energy offset.

Before discussing specific issues related to the calculation of the force, we deem useful to give a brief summary of the DMC algorithm. The Green function is approximated, with a bias that vanishes in the limit of zero time step, by an explicit form such as

$$\mathcal{G}(R', R) = \frac{\exp\{-[R' - R - V(R)\tau]^2/2\tau\}\exp[S(R', R)]}{(2\pi\tau)^{3N/2}} \tag{5}$$

with $V(R) = \nabla\Psi(R)/\Psi(R)$ and $S(R',R) = \tau\{E_T - [E_L(R') + E_L(R)]/2\}$ (in the actual simulation we use the improved $\bar{V}$ and $\bar{S}$ introduced in ref 28). The simulation proceeds through a branching random walk of a population of walkers. A walker at R drifts to $R + V(R)\tau$ and then diffuses to $\tilde{R}$ according to the drift-diffusion part $T(\tilde{R},R) = \exp\{-[\tilde{R} - R - V(R)\tau]^2/2\tau\}$ of the Green function of eq 5. The new configuration R' of the walker is either $\tilde{R}$ or R with probability

$$p = \min\left\{1, \frac{|\Psi(\tilde{R})|^2 T(R, \tilde{R})}{|\Psi(R)|^2 T(\tilde{R}, R)}\right\} \tag{6}$$

or $1 - p$, respectively. Finally, the statistical weight of the walker is multiplied by $\exp[S(R',R)]$. Walkers with large (small) weight are replicated (deleted) to control the fluctuations of individual weights, and the offset E_T is adjusted to control the fluctuation of the total weight of the population. In the limit of perfect importance sampling (i.e., $\Psi = \Phi$), the branching term $\exp[S(R',R)]$ can be set to unity, and the rejection step of eq 6 ensures that the distribution Ψ^2 is sampled exactly, despite the short-time approximation in the Green function, by imposing detailed balance. In general, the rejection step is instrumental to achieve small time step errors.

We now address the calculation of the force. After the random walk has reached equilibrium, we can write $\tilde{P}$ as

$$\tilde{P}(R_n) = \int \mathrm{d}R_{n-1}\ldots\mathrm{d}R_{n-k}\Big[\prod_{i=n-k}^{n-1}\mathcal{G}(R_{i+1}, R_i)\Big]\tilde{P}(R_{n-k}) \tag{7}$$

where n is the current iteration. This expression allows for explicit differentiation of the DMC energy. For, replacing $P(R_n) = \tilde{P}(R_n)/\int \mathrm{d}R\tilde{P}(R)$ in eq 1, we obtain

$$\begin{aligned} E &= \int \mathrm{d}R_n\ldots\mathrm{d}R_{n-k}E_L(R_n)\Big[\prod_{i=n-k}^{n-1}\mathcal{G}(R_{i+1}, R_i)\Big]\tilde{P}(R_{n-k}) \\ &\quad \Big/\int \mathrm{d}R_n\ldots\mathrm{d}R_{n-k}\Big[\prod_{i=n-k}^{n-1}\mathcal{G}(R_{i+1}, R_i)\Big]\tilde{P}(R_{n-k}) \\ &\equiv \int \mathrm{d}R_n\ldots\mathrm{d}R_{n-k}E_L(R_n)\Pi(R_{n-k}, \ldots, R_n) \end{aligned} \tag{8}$$

and differentiation with respect to the nuclear coordinates leads to

$$\begin{aligned} F = -\langle &\nabla_\alpha E_L(R_n) + [E_L(R_n) - E] \\ &\times \Big[\sum_{i=n-k}^{n-1}\nabla_\alpha \ln \mathcal{G}(R_{i+1}, R_i) + \nabla_\alpha \ln \tilde{P}(R_{n-k})\Big]\rangle_\Pi \end{aligned} \tag{9}$$

where $\langle f(R_{n-k},\ldots,R_n)\rangle_\Pi$ denotes the integral over the probability distribution $\Pi(R_{n-k},\ldots,R_n)$ defined in eq 8, namely, over the paths generated in the drift-diffusion-branching random walk. The unknown last term in square brackets does not contribute to the average provided $k\tau$ is larger than the correlation time between $E_L(R_n)$ and $\tilde{P}(R_{n-k})$, because in this case

$$\begin{aligned} \langle [E_L(R_n) - E]\nabla_\alpha \ln \tilde{P}(R_{n-k})\rangle_\Pi &= \langle E_L(R_n) - E\rangle_\Pi \\ &\times \langle \nabla_\alpha \ln \tilde{P}(R_{n-k})\rangle_\Pi = 0 \end{aligned} \tag{10}$$

Replacing eq 5 in eq 9, we see that the calculation of the nonvanishing terms of the force involves only derivatives of the trial function and of the local energy, just as in VMC. Technically, all that DMC adds is the minor complication of storing the relevant quantities not only for the current configuration but also for the last k steps of the random walk.

However, there are three more issues that deserve further analysis, namely, (i) the accept/reject step and the noise in the drift-diffusion term, (ii) the divergence of the variance, and (iii) the optimization of the trial function.

(i). In the calculation of $\nabla_\alpha \ln \mathcal{G}$ in eq 9, when the move is rejected, we naively use $\mathcal{G}(R, R) = \exp\{-[V(R)\tau]^2/2\tau\}\exp[S(R, R)]$. This would correspond to a drift from R to $R+V(R)\tau$ followed by a diffusion back to R, which is not quite what the simulation does of diffusing to $\tilde{R}$ and rejecting the move with probability $1 - p$. As a result, we find a strong time step dependence of the force (see the solid line in Figure 1). This problem could presumably be mitigated, perhaps calculating $\nabla_\alpha \ln[pT(\tilde{R},R) + (1 - p) \times 1]$, but we have not investigated this issue any further because $\nabla_\alpha \ln T$ has large statistical fluctuations that would make the full estimator (eq 9) rather inefficient even if the time step dependence were not as bad.

(ii). In both VMC and DMC, the estimator of the force (eq 2) has zero variance if the trial function is exact but a diverging variance for an approximate trial function. In DMC, the divergence of the variance of F can be removed by performing the accept/reject step (eq 6) with the square of a modified trial function $\tilde{\Psi}_\tau(R)$, which is finite at the nodes and reduces to Ψ in the limit of small τ. The bias incurred vanishes at zero time step. The averages are then reweighted with the factor $\Psi^2(R)/\tilde{\Psi}_\tau^2(R)$ to ensure their correctness even for finite τ in the limit of perfect importance sampling. We choose here the same functional form of $\tilde{\Psi}_\tau$ as in ref 12, but scale by $\sqrt{\tau}$

dx.doi.org/10.1021/ct500780r | J. Chem. Theory Comput. 2014, 10, 4823−4829

the distance from the nodes over which $\tilde{\Psi}_\tau$ differs from Ψ.

(iii). In FN-DMC, the dependence of the energy on the nuclear coordinates through the variational parameters is restricted to those parameters that directly change the nodal surface; on general grounds, this dependence is expected to be milder than in VMC. However, we optimize the variational parameters (as almost invariably done) by minimizing the VMC energy rather than the FN energy, so that, in general, the DMC force is biased by the neglect of the unknown term shown in eq 3. For the sake of verifying that the full force, eq 9, is consistent with the gradient of the PES, in Figure 1 we have calculated the PES with orbitals independent of the atomic distance, $\nabla_\alpha c_i = 0$.

To summarize, this section presents a procedure to calculate the forces in FN-DMC with finite variance. In the academic but not trivial case of orbitals independent of the atomic position, the force is unbiased, as shown by the solid line in Figure 1. Unfortunately, such a procedure seems of little practical interest because of the large statistical noise coming from the drift-diffusion term, let alone the problematic time step dependence of the present implementation.

2.3. The Variational Drift-Diffusion Approximation. To overcome these difficulties, an approximate DMC method for the computation of interatomic forces, which we dub Variational Drift-Diffusion (VD), was introduced in refs 15 and 29, in the framework of finite increments and correlated sampling. In computing the force as energy difference for different geometries, we distinguish between a primary and a secondary walk for the reference and a secondary geometry, with nuclear coordinates R_α and R^s_α, trial wave functions Ψ and Ψ_s, and corresponding local energies E_L and E^s_L, respectively. The secondary walk is constructed from the coordinates of the primary walk, rather than independently simulated, and its properties are calculated via correlated sampling. The method is devised to be as close as possible to VMC and to exactly recover the DMC algorithm with accept/reject for the primary geometry, while yielding results very close to the DMC value for the secondary geometry.

The starting observation is that, if we neglect the growth/decay factor, a DMC algorithm in the presence of the accept/reject step reduces to the VMC algorithm. Therefore, it results in the sampling of Ψ^2 for the primary walk and Ψ^2_s for the secondary walk as long as we reweight the averages for the secondary geometry with the ratio of the square of the secondary and primary wave functions. Formally, we recover the DMC solution by multiplying the weights of the primary and secondary walkers by the corresponding growth/decay factors, in the sense of the Pure Diffusion Monte Carlo method.[30] This gives the exact distribution for the primary walk, but only an approximate distribution for the secondary walk, since the moves were not proposed with the dynamics appropriate to the secondary geometry, namely, with the drift-diffusion part of the Green function computed for the secondary wave function and the corresponding accept/reject probability. Therefore, the DMC energy difference is approximated as

$$E^s - E = \frac{\langle E^s_L(R_n)\Delta W^s_n\rangle_\Pi}{\langle \Delta W^s_n\rangle_\Pi} - \langle E_L(R_n)\rangle_\Pi \tag{11}$$

where the averages are over the mixed distribution $\Psi\Phi$ of the primary walk and the secondary energy contains the weights

$$\Delta W^s_n = \frac{|\Psi_s(R_n)|^2}{|\Psi(R_n)|^2}\prod_{i=n-k}^{n-1}\exp[\Delta S(R_{i+1}, R_i)] \tag{12}$$

with $\Delta S = S^s - S$. Finally, if the space-warp transformation is employed to generate the secondary walk,[15,27] only the Jacobian of the transformation $J(R_n)$ of the current configuration is included in ΔW^S_n, rather than one Jacobian for each step of the random walk involved in eq 12. This choice is consistent with the main approximation of the algorithm: both are designed to yield the correct energy for the secondary geometry in the limit of perfect importance sampling.

It is of course possible to formulate a differential expression of the finite energy difference (11), which gives the analytical force:

$$F = -\langle \nabla_\alpha E_L(R_n) + [E_L(R_n) - E] \times [2\nabla_\alpha \ln\Psi(R_n) + \sum_{i=n-k}^{n-1}\nabla_\alpha S(R_{i+1}, R_i)]\rangle_\Pi \tag{13}$$

If one uses the space-warp transformation, an additional term $\nabla_\alpha \ln J(R_n)$ must be included in the last bracket.

In short, the VD approximation[15] replaces the logarithmic derivative of the drift-diffusion part of the Green function along the random walk in eq 9 with the logarithmic derivative of the trial function at the current configuration, eq 13; these two quantities have the same average in the limit of perfect importance sampling and zero time step since the drift-diffusion-branching equation reduces then to the Fokker–Planck equation with asymptotic probability density given by the square of the trial wave function. The VD algorithm avoids the problems of item (i) of the previous section; it affords a roughly 10-fold increase in efficiency for given time step and features a much better behavior of the time step dependence; see the dotted line in Figure 1. The bias introduced depends on the quality of the trial function. It is generally very small, as we will show in section 4. The large discrepancy between the approximate force and the gradient of the PES shown in Figure 1 is due to an intentionally poor trial function (a one-parameter, two-body Jastrow factor and HF orbitals independent of the atomic distance).

3. DETAILS OF THE SIMULATIONS

We use the program package CHAMP[31] with a small cc-pVDZ basis set and scalar-relativistic energy-consistent Hartree–Fock pseudopotentials.[32,33] With the only exception of the calculations in Figure 1, the pseudopotentials are treated in DMC using the "T-moves" approximation of ref 34, an upgrade of the standard locality approximation,[35] and the forces are calculated using the space-warp transformation. The trial function is the product of a Jastrow factor[36] and an antisymmetric factor consisting of either a single determinant or a linear combination thereof. The Jastrow factor is either limited to two-body electron–electron and electron–nucleus correlations, or includes three-body electron–electron–nucleus correlations. The determinantal component of the trial function—Hartree–Fock, DFT-B3LYP, or complete-active-space (CAS) expansion—is initially obtained with the program GAMESS (US).[37] The one-particle orbitals and/or the configuration interaction (CI) coefficients of the CAS expansion are then optionally optimized in VMC using the

linear method.[5,6] For C_2, the CAS expansion correlates 6 electrons in the 6 orbitals ($3\sigma_g$, $1\pi_u^x$, $1\pi_u^y$, $1\pi_g^x$, $1\pi_g^y$, $3\sigma_u$). Table 1 specifies the various trial functions used in this work.

Table 1. Name and Specification of the Trial Functions[a]

name	Jastrow	determinant(s)
HF_{1pJ2}	2-body, 1 par.	1-det, HF orbitals
HF_{J3}	3-body	1-det, HF orbitals
$B3LYP_{J2}$	2-body	1-det, B3LYP orbitals
$B3LYP_{J3}$	3-body	1-det, B3LYP orbitals
$1DET_{all}$	3-body	1-det, optimized orbitals
CAS_{CI}	3-body	CAS, optimized CI coeff.
CAS_{all}	3-body	CAS, optimized CI coeff. and orbitals

[a]For HF_{1pJ2}, the HF orbitals and the optimal Jastrow parameter are calculated for $d = 2.4$ and held fixed at other distances.

4. RESULTS

We calculate the forces for the SiH and C_2 molecules on a range of the atomic distance d around the equilibrium bond length d_0. SiH is chosen in order to compare the VD approximation with the "total pure force" F^{tot}_{pure} of ref 19, which we consider as a high standard of accuracy. C_2, a more challenging system with a strong multiconfiguration character, is chosen in order to test the VD approximation on wave functions of widely different quality. The central quantity of interest is the error Δ, defined as the difference between the force obtained from eq 2 and the derivative of the PES obtained as a fit to the QMC energies. We recall that the error stems from DMC-specific approximations such as the VD estimator, eq 13, or the total pure (TP) force, eq 21 of ref 19, and from neglect of the term in eq 3 whenever relevant (in DMC, when Ψ is optimized by minimization of the VMC energy, and in VMC when the optimization is not complete). Hartree atomic units are used throughout.

Figure 3 shows the error of the force for the SiH molecule using an HF_{J3} trial function. The DMC-VD error (circles) is close to 0.001, with a weak dependence on d in the range considered. The DMC-TP errors (solid lines, taken from Figure 2 of ref 19) have a similar magnitude on average but a stronger dependence on d which might be of concern for very compressed or very stretched bonds. The TP approximation

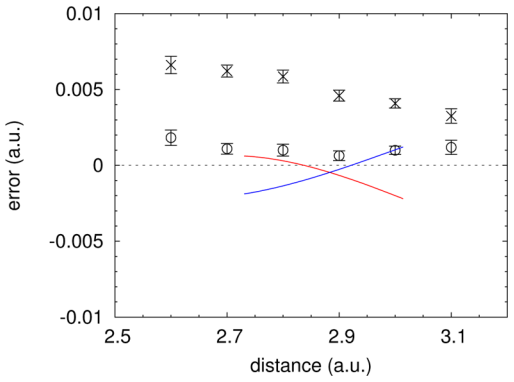


Figure 3. Error on the force Δ as a function of the distance d for the SiH molecule, using an HF_{J3} trial function. × VMC; ○ DMC in the VD approximation. Solid lines: F^{tot}_{pure} from ref 19 calculated on the H atom (red) and the Si atom (blue). The statistical error on Δ mainly comes from the fit to the QMC energies.[25]

also features a peculiar difference between the force calculated on the two atoms. The comparison illustrated in Figure 3 strongly supports the accuracy of the DMC-VD forces. Note that the VMC forces have a significantly larger error (crosses), due to the presence of HF orbitals rather than optimized orbitals in the trial function.

Hereafter, all DMC forces are understood to be calculated with the VD algorithm. Table 2 collects several results for C_2 obtained in VMC and DMC using different trial functions specified in section 3. Obviously, better trial functions give lower energy; the entry E_{min} shows that (i) DFT orbitals with the B3LYP functional are better than HF (in VMC for given Jastrow factor, and in DMC regardless of the Jastrow factor), and nearly as good as optimized orbitals, and (ii) CAS wave functions are much superior (indeed, the equilibrium bond length d_0 is in agreement with the experimental value only in DMC with CAS wave functions). However, the primary interest here is the consistency between the force and the PES, measured by the difference $d_0^F - d_0^E$ between the distances of zero force and minimum energy, as well as by the error Δ itself. The former measure relates to the criterion for good accuracy proposed in ref 19, namely, $d_0^F - d_0^E$ should be significantly smaller than 0.02, the typical error of DFT on d_0, for a QMC calculation of forces to be useful. The latter is a more global assessment of accuracy, because it involves a range of atomic distances. The entry $\overline{|\Delta|}$ of Table 2 is the unsigned error averaged over distances between 2.1 and 2.7, and the Δ as a function of d is shown in Figures 4−6 for several cases.

In VMC, the force is exact, within statistical uncertainties, for the trial functions $1DET_{all}$ and CAS_{all} because of full optimization. Incomplete optimization causes very large errors with HF orbitals; see the squares in Figure 4; with B3LYP orbitals (circles and downward triangles in Figure 4), as well as with CAS_{CI} (crosses in Figure 5), the errors are much smaller, $|\Delta| \lesssim 0.01$ but not completely negligible. In particular, the VMC values of $d_0^F - d_0^E$ with the B3LYP trial functions are extremely small just because Δd happens to cross zero for $d \simeq d_0$.

Comparison of Figures 4 and 5 with Figure 6, and of the VMC and DMC results in Table 2 shows that, whenever the VMC force is approximated, the errors are smaller in DMC than in VMC. Upon full optimization, the VMC errors drop to zero (within the statistical noise), while the DMC errors change rather mildly without a clear trend: the HF_{J2} trial function is the only instance with a distinct correlation between quality of Ψ and accuracy of the DMC force; both are poor in this case, but the accuracy is still better than that in VMC. HF orbitals aside, all the DMC forces give a discrepancy $d_0^F - d_0^E$ well below the "usefulness threshold"[19] 0.02 (see Table 2), and an error $|\Delta| < 0.005$ with a weak dependence on d (see Figure 6). While it may be somewhat disappointing that the error is not smallest for the best trial function, the general accuracy and robustness of the DMC force is very satisfying.

We finally discuss the efficiency of the DMC calculation of forces in the VD approximation. The DMC algorithm requires a small time step, therefore the calculation of all quantities is less efficient than in VMC. Furthermore, for the VD force, there are the additional terms ∇S in eq 13, not present in the VMC force, which are expected to increase the variance. Figure 7 shows the ratio between the statistical error on the force σ_F and the statistical error on the energy σ_E calculated in the same run for the carbon dimer. In VMC, this ratio varies between 1 and 2.5 in the range of distances and with the trial function

dx.doi.org/10.1021/ct500780r | J. Chem. Theory Comput. 2014, 10, 4823−4829

Table 2. VMC and DMC Results for the Carbon Dimer Using Different Trial Functions[a]

Ψ	d_0^E	d_0^F	$d_0^F - d_0^E$	$\overline{\lvert\Delta\rvert}$	E_{min}
			VMC		
HF_{J3}	2.3677(12)	2.4200(2)	0.0522(12)	0.0373(12)	−11.01269(16)
$B3LYP_{J2}$	2.3915(11)	2.3883(2)	−0.0032(11)	0.0050(10)	−11.00213(13)
$B3LYP_{J3}$	2.3783(12)	2.3799(2)	0.0016(12)	0.0051(12)	−11.02198(17)
$1DET_{all}$	2.3785(12)	2.3774(2)	−0.0011(12)	0.0009(12)	−11.02430(15)
CAS_{CI}	2.3592(7)	2.3532(2)	−0.0059(7)	0.0046(8)	−11.07351(15)
CAS_{all}	2.3548(7)	2.3542(4)	−0.0006(8)	0.0006(11)	−11.07556(9)
			DMC		
HF_{J3}	2.3724(12)	2.4089(4)	0.0365(12)	0.0266(14)	−11.04696(14)
$B3LYP_{J2}$	2.3799(14)	2.3839(3)	0.0040(14)	0.0027(14)	−11.05048(17)
$B3LYP_{J3}$	2.3763(11)	2.3798(3)	0.0035(11)	0.0020(12)	−11.05348(14)
$1DET_{all}$	2.3755(11)	2.3751(3)	−0.0004(11)	0.0015(11)	−11.05413(13)
CAS_{CI}	2.3494(7)	2.3499(2)	0.0005(7)	0.0008(8)	−11.08759(9)
CAS_{all}	2.3483(6)	2.3523(2)	0.0040(7)	0.0029(8)	−11.08860(9)

[a]Equilibrium bond length d_0^E obtained from the PES;[25] equilibrium bond length d_0^F obtained from the force;[25] error $d_0^F - d_0^E$ on d_0; mean unsigned error on the force $\overline{\lvert\Delta\rvert}$; minimum energy E_{min}. The statistical error shown on $\overline{\lvert\Delta\rvert}$ is the average over d of the statistical error on Δ (Figures 4 and 5 for VMC and Figure 6 for DMC). The experimental value of the equilibrium bond length is[38] $d_0^{EXP} = 2.348$.

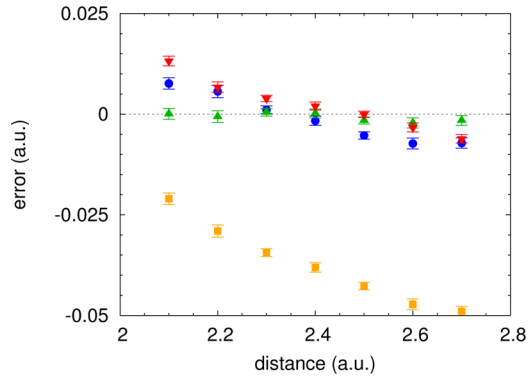


Figure 4. Error on the force Δ as a function of the distance d for the C_2 molecule, computed in VMC with a 1-determinant wave function. ▲ (green) $1DET_{all}$; ● (blue) $B3LYP_{J3}$; ▼ (red) $B3LYP_{J2}$; ■ (orange) HF_{J3}.

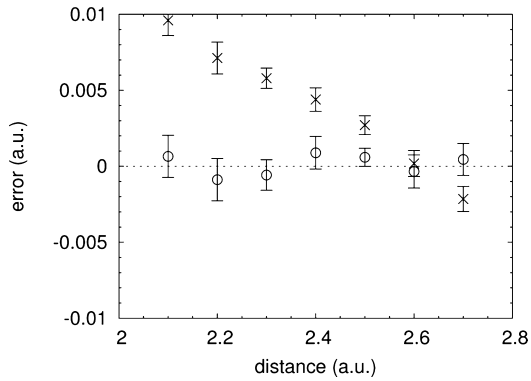


Figure 5. Error on the force Δ as a function of the distance d for the C_2 molecule, computed in VMC with a CAS wave function. ○ CAS_{all}; × CAS_{CI}.

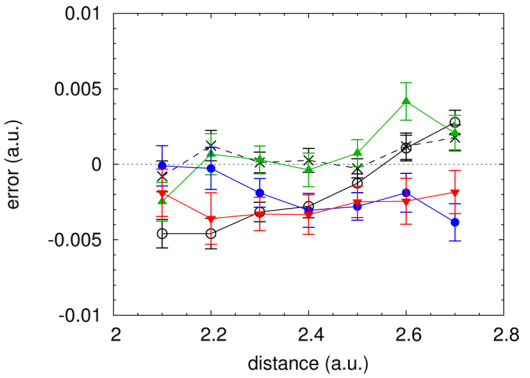


Figure 6. Error on the force Δ as a function of the distance d for the C_2 molecule, computed in DMC with various wave functions. ○ CAS_{all}; × CAS_{CI}; ▲ (green) $1DET_{all}$; ● (blue) $B3LYP_{J3}$; ▼ (red) $B3LYP_{J2}$.

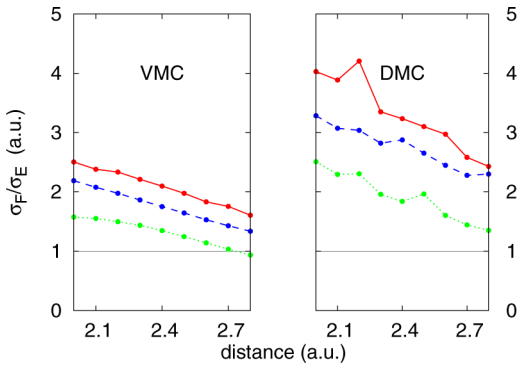


Figure 7. Ratio of the statistical errors on the force σ_F and on the energy σ_E computed in the same run. We show VMC (left panel) and DMC (right panel) results obtained using different trial functions: solid line, CAS_{all}; dashed line, $1DET_{all}$; dotted line, HF_{J3}.

considered (note that the ratio is larger the better the trial function: this does not mean that the calculation of the force becomes more expensive, but rather that the energy has a stronger zero-variance property than the force). In DMC, we have a very similar behavior, scaled up by a factor of roughly 1.6. This implies that if the energy is α times more expensive in DMC than in VMC, the force is $\sim 1.6^2\alpha$ more expensive in DMC than in VMC. We conclude that in DMC the calculation of the VD force is only slightly less efficient than the calculation of the energy.

dx.doi.org/10.1021/ct500780r | J. Chem. Theory Comput. 2014, 10, 4823−4829

5. CONCLUSIONS

We have investigated the performance of the approximated VD method for the computation of interatomic forces in DMC, with the intent to establish whether it represents a practical scheme leading to higher accuracy than the use of VMC forces. We have tested the approach for the SiH and C_2 molecules by computing the forces and the derivatives of the PES in VMC and DMC over a range of atomic distances to determine the degree of consistency between forces and energy. For C_2, we have employed a variety of wave functions optimized in VMC to study the dependence of the errors on the forces and the resulting equilibrium bond lengths on the quality of the trial functions.

For SiH, a comparison of the VD results with those obtained with the approximate "total pure force" DMC approach[19] reveals that both methods lead to equally small discrepancies between forces and derivatives of the PES near the equilibrium bond length but that the errors in the DMC-VD scheme display a significantly weaker dependence on the interatomic distance. The calculations for the more challenging C_2 molecule demonstrate that the choice of trial function has a rather small influence on the accuracy of the VD approximation with wave functions optimized in VMC: For all but the poorest wave function, the errors in the DMC forces are well below 0.005 au and the equilibrium bond lengths obtained from the DMC force and the PES consistently agree to better than 0.004 au Importantly, whenever the VMC forces are approximate due to a partial optimization of the trial function, the DMC errors are consistently smaller than the VMC ones. The best results for the equilibrium bond length calculated from the force are in fact obtained at the DMC level.

Finally, a crucial feature of the VD scheme is its computational efficiency, since the calculation of interatomic DMC-VD forces has a cost only slightly higher than the computation of the DMC energy. We can therefore conclude that the VD approximation is a robust method for the computation of accurate DMC forces, offering a cost-effective route to systematically improve on the use of VMC forces.

AUTHOR INFORMATION

Corresponding Authors

*Email: moroni@democritos.it.
*Email: c.filippi@utwente.nl.

Notes

The authors declare no competing financial interest.

REFERENCES

(1) Foulkes, W. M. C.; Mitas, L.; Needs, R. J.; Rajagopal, G. *Rev. Mod. Phys.* **2001**, *73*, 33.
(2) Kolorenč, J.; Mitas, L. *Rep. Prog. Phys.* **2011**, *74*, 026502.
(3) Wagner, L. *Int. J. Quantum Chem.* **2014**, *114*, 94.
(4) We adopt the convention that the cost of the force is the same as the cost of the energy if the statistical errors on the force and on the energy per particle, expressed in au and calculated in the same run, are equal.
(5) Umrigar, C. J.; Toulouse, J.; Filippi, C.; Sorella, S.; Hennig, R. G. *Phys. Rev. Lett.* **2007**, *98*, 110201.
(6) Toulouse, J.; Umrigar, C. J. *J. Chem. Phys.* **2007**, *126*, 084102.
(7) Neuscamman, E.; Umrigar, C. J.; Chan, G. K.-L. *Phys. Rev. B* **2012**, *85*, 045103.
(8) Valsson, O.; Filippi, C. *J. Chem. Theory Comput.* **2010**, *6*, 1275.
(9) Guareschi, R.; Filippi, C. *J. Chem. Theory Comput.* **2013**, *9*, 5513.
(10) Saccani, S.; Filippi, C.; Moroni, S. *J. Chem. Phys.* **2013**, *138*, 084109.
(11) Barborini, M.; Guidoni, L. *J. Chem. Phys.* **2012**, *137*, 224309.
(12) Attaccalite, C.; Sorella, S. *Phys. Rev. Lett.* **2008**, *100*, 114501.
(13) Assaraf, R.; Caffarel, M.; Kollias, A. C. *Phys. Rev. Lett.* **2011**, *106*, 150601.
(14) Wagner, L. K.; Grossman, J. C. *Phys. Rev. Lett.* **2010**, *104*, 210201.
(15) Filippi, C.; Umrigar, C. J. *Phys. Rev. B* **2000**, *61*, R16291.
(16) Barnett, R. N.; Reynolds, P. J.; Lester, W. A. *J. Comput. Phys.* **1991**, *96*, 258.
(17) Casalegno, M.; Mella, M.; Rappe, A. M. *J. Chem. Phys.* **2003**, *118*, 7193.
(18) Lee, M. W.; Mella, M.; Rappe, A. M. *J. Chem. Phys.* **2005**, *122*, 244103.
(19) Badinski, A.; Haynes, P. D.; Trail, J. R.; Needs, R. J. *J. Phys.: Condens. Matter* **2010**, *22*, 074202.
(20) Assaraf, R.; Caffarel, M. *J. Chem. Phys.* **2000**, *113*, 4028.
(21) Assaraf, R.; Caffarel, M. *J. Chem. Phys.* **2003**, *119*, 10536.
(22) Chiesa, S.; Ceperley, D. M.; Zhang, S. *Phys. Rev. Lett.* **2005**, *94*, 036404.
(23) Liu, K. S.; Kalos, M. H. *Phys. Rev. A* **1974**, *10*, 303.
(24) Sorella, S.; Capriotti, L. *J. Chem. Phys.* **2010**, *133*, 234111.
(25) We fit a polynomial to a given set of QMC values for either the energy or the force. We then produce a large number of fake data sets normally distributed around the fitted polynomial, with the QMC statistical errors as standard deviation. Finally, we fit a polynomial to each set of fake data. All derived quantities and their statistical errors, such as the force or the equilibrium bond length, are obtained as averages over the fake data sets and their standard deviations.
(26) Badinski, A.; Needs, R. J. *Phys. Rev. B* **2008**, *78*, 035134.
(27) Umrigar, C. J. *Int. J. Quantum Chem.* **1989**, 217−230.
(28) Umrigar, C. J.; Nightingale, M. P.; Runge, K. J. *J. Chem. Phys.* **1993**, *99*, 2865.
(29) Filippi, C.; Umrigar, C. J. Interatomic Forces and Correlated Sampling in Quantum Monte Carlo. *Recent Advances in Quantum Monte Carlo Methods, Part II*; World Scientific: Singapore, 2002.
(30) Caffarel, M.; Claverie, P. *J. Chem. Phys.* **1988**, *88*, 1088.
(31) CHAMP is a quantum Monte Carlo program package written by C. J. Umrigar, C. Filippi, and collaborators.
(32) Burkatzki, M.; Filippi, C.; Dolg, M. *J. Chem. Phys.* **2007**, *126*, 234105.
(33) For the hydrogen atom, we use a more accurate BFD pseudopotential and basis set. Dolg, M.; Filippi, C., private communication.
(34) Casula, M. *Phys. Rev. B* **2006**, *74*, 161102.
(35) Mitas, L.; Shirley, E. L.; Ceperley, D. M. *J. Chem. Phys.* **1991**, *95*, 3467.
(36) Filippi, C.; Umrigar, C. J. *J. Chem. Phys.* **1996**, *105*, 213−226 As Jastrow correlation factor, we use the exponential of the sum of three fifth-order polynomials of the electron−nuclear (e-n), the electron−electron (e-e), and of pure 3-body mixed e-e and e-n distances (e-e-n), respectively. A 2-body Jastrow factor only includes e-n and e-n terms. The Jastrow factor is adapted to deal with pseudo-atoms, and the scaling factor κ is set to 0.6 au.
(37) Schmidt, M. W.; Baldridge, K. K.; Boatz, J. A.; Elbert, S. T.; Gordon, M. S.; Jensen, J. H.; Koseki, S.; Matsunaga, N.; Nguyen, K. A.; Su, S.; Windus, T. L.; Dupuis, M.; Montgomery, J. A., Jr. *J. Comput. Chem.* **1993**, *14*, 1347−1363.
(38) Huber, K.; Herzberg, G. In *Molecular Spectra and Molecular Structure. IV. Constants of Diatomic Molecules*; Van Nostrand Reinhold Company: New York, 1979.