

INTERATOMIC FORCES AND CORRELATED SAMPLING IN QUANTUM MONTE CARLO

CLAUDIA FILIPPI

Department of Physics, National University of Ireland, Cork, Ireland

C. J. UMRIGAR

*Cornell Theory Center and Laboratory of Atomic and Solid State Physics,
Cornell University, Ithaca, New York 14853*

One of the main difficulties of quantum Monte Carlo techniques is the lack of an efficient method for computing interatomic forces. To date, most quantum Monte Carlo calculations have been performed on geometries obtained with either density functional theory or conventional quantum chemistry methods. Here, we present a correlated sampling method to efficiently calculate numerical forces and potential energy surfaces in diffusion Monte Carlo. It employs a novel coordinate transformation, earlier used in variational Monte Carlo, to greatly reduce the statistical error. Results are presented from all-electron and pseudopotential calculations of homonuclear diatomic molecules.

1 Introduction

Over the past decade, quantum Monte Carlo (QMC) methods [1,2] have provided the most accurate calculations of correlated properties for medium-sized molecules and solid systems, where conventional quantum chemistry methods become computationally very expensive. Despite their potential to become an important tool for the study of correlated systems, a major stumbling block to a more widespread use of QMC techniques has been their difficulty in determining equilibrium geometries and potential energy surfaces: most QMC calculations have been performed on geometries obtained with either density functional theory (DFT) or conventional quantum chemistry methods.

To understand the complications involved in computing forces in QMC, we need to first discuss how other conventional electronic structure methods determine interatomic forces, and why such procedures cannot be straightforwardly extended to QMC calculations. DFT methods or standard quantum chemistry techniques use the Hellman-Feynman theorem to compute forces on nuclei [3]. The Hellman-Feynman theorem is applicable provided that the parameters in the wave function are chosen to minimize the energy for the given form of the wave function (Pulay [4] corrections must be added if the basis employed depends on the atomic coordinates and is not complete). The applicability of the Hellman-Feynman theorem in QMC methods is compli-

cated by the following considerations. First, the wave functions used in QMC are usually not obtained by minimizing the energy, but rather by minimizing the variance of the local energy. Although these conditions are identical if no constraints are placed on the form of the wave function, for the wave functions used in practice they are close but not identical. Therefore, if the Hellman-Feynman theorem were employed in variational Monte Carlo (VMC), the forces would have a systematic error. (When we speak of the error in the force we do not mean the error compared to the true force but rather the error relative to the force obtained by calculating the energy at neighboring geometries using the same form of the wave function with reoptimized parameters.) At first sight, it may appear that the systematic error disappears in diffusion Monte Carlo (DMC) since this method stochastically projects onto the true ground state. However, practical DMC calculations require one to employ the fixed-node approximation [2] to avoid an exponential growth of the bosonic ground state relative to the fermionic ground state as the projection time increases. For fixed-node DMC it has been shown [5] that the Hellman-Feynman theorem is applicable provided that at least one of the following two conditions is satisfied: a) the trial wave function has the true nodes or b) the nodes of the trial functions are independent of the nuclear positions. In fact, the former condition can be relaxed to read: the parameters of the trial wave function have been optimized to give the lowest fixed-node DMC energy for that form of the wave function, and the basis used to construct the wave function is either not dependent on the atomic positions or is complete. The other problem with using the Hellman-Feynman theorem in both VMC and DMC is that a straightforward application of the theorem leads to an infinite variance estimator for the force, thereby rendering the method totally impractical. However, a recently developed and general variance reduction method, due to Assaraf and Caffarel [6], makes the variance finite and small.

Alternatively, one could simply compute energy differences to obtain either forces (for an infinitesimal displacement of the ions) or the full potential energy surface of the system. However, while quantum chemistry methods can rely on having an approximately constant and smoothly varying error in the energy, a major disadvantage of QMC methods is that, in addition to systematic errors, one has statistical errors which make the determination of energy differences or smooth potential energy surfaces computationally very expensive. Even though it is not possible to entirely eliminate the statistical errors, it is possible, by using correlated sampling [2], to make the statistical errors in the relative energies of different geometries much smaller than the errors in the separate energies and to make them vanish in the limit that the two

geometries become identical. In the past, the correlated sampling technique has been used within VMC [7,8] but there have been very few attempts [9] to extend the approach to DMC, and these were approximate and/or inefficient and were tested only on H_2 , H_3^+ and LiH .

In this paper, we discuss our DMC correlated sampling technique [10] for computing accurate forces, potential energy surfaces and vibration frequencies. The DMC bond lengths of first-row diatomic molecules computed with this algorithm are found to be in better agreement with experimental values than are the VMC, Hartree-Fock (HF), local density approximation (LDA) and generalized gradient approximation (GGA) values.

2 Correlated sampling in variational Monte Carlo

Correlated sampling enables us to compute, from a single reference Monte Carlo walk, the relative energies of different geometries, a reference and one or more secondary geometries, with nuclear coordinates \mathbf{R}_α and \mathbf{R}_α^s , Hamiltonians \mathcal{H} and \mathcal{H}_s , and wave functions ψ and ψ_s , respectively. Unbiased expectation values are obtained by reweighting the configurations sampled from ψ^2 ,

$$\begin{aligned} E_s - E &= \frac{\langle \psi_s | \mathcal{H}_s | \psi_s \rangle}{\langle \psi_s | \psi_s \rangle} - \frac{\langle \psi | \mathcal{H} | \psi \rangle}{\langle \psi | \psi \rangle} \\ &= \frac{1}{N_{\text{conf}}} \sum_{i=1}^{N_{\text{conf}}} \left\{ \frac{\mathcal{H}_s \psi_s(\mathbf{R}_i)}{\psi_s(\mathbf{R}_i)} W_i - \frac{\mathcal{H} \psi(\mathbf{R}_i)}{\psi(\mathbf{R}_i)} \right\}, \end{aligned} \quad (1)$$

where the weights of the N_{conf} MC configurations are

$$W_i = \frac{N_{\text{conf}} |\psi_s(\mathbf{R}_i)/\psi(\mathbf{R}_i)|^2}{\sum_{i=1}^{N_{\text{conf}}} |\psi_s(\mathbf{R}_i)/\psi(\mathbf{R}_i)|^2}, \quad (2)$$

and $\mathbf{R} \equiv (\mathbf{r}_1, \dots, \mathbf{r}_N)$. The effective number of configurations sampled is

$$N_{\text{eff}} = \frac{\left(\sum_{i=1}^{N_{\text{conf}}} W_i \right)^2}{\sum_{i=1}^{N_{\text{conf}}} W_i^2} \quad (3)$$

so that a gain in efficiency is achieved so long as the weights do not fluctuate wildly. The statistical error in $E_s - E$ goes to zero linearly as the secondary geometry approaches the reference geometry, thereby permitting the evaluation of numerical forces with a finite statistical error, that can be reduced by increasing the computer time.

2.1 Space-warp coordinate transformation

The electronic coordinates sampled from the reference wave function squared, ψ^2 , will not be optimal for computing the energy E_s corresponding to the nuclear coordinates \mathbf{R}_α^s , since the electron density will be peaked at \mathbf{R}_α rather than at \mathbf{R}_α^s . This problem can be solved by mapping the electron coordinates so that the electrons close to a given nucleus move almost rigidly with that nucleus [8]:

$$\mathbf{r}_i^s = \mathbf{r}_i + \sum_{\alpha=1}^{N_{\text{atoms}}} (\mathbf{r}_\alpha^s - \mathbf{r}_\alpha) \omega_\alpha(\mathbf{r}_i), \quad (4)$$

where

$$\omega_\alpha(\mathbf{r}_i) = \frac{F(|\mathbf{r}_i - \mathbf{r}_\alpha|)}{\sum_{\beta=1}^{N_{\text{atoms}}} F(|\mathbf{r}_i - \mathbf{r}_\beta|)}; \quad \sum_{\alpha=1}^{N_{\text{atoms}}} \omega_\alpha(\mathbf{r}_i) = 1. \quad (5)$$

(We use Latin indices for electronic coordinates and Greek indices for nuclear coordinates.) $F(r)$ is any sufficiently rapidly decaying function, e.g., $r^{-\kappa}$, $e^{-\kappa r}$ or $e^{\kappa/r}$. The reduction in statistical error is large (see Fig. 2) and almost independent of the choice for $F(r)$. In this paper, we make the same choice as in our earlier work [8], $F(r) = r^{-\kappa}$ and $\kappa = 4$, but other choices yield almost the same reduction in statistical error.

The equation for $E_s - E$ (Eq. 1) is now

$$E_s - E = \frac{1}{N_{\text{conf}}} \sum_{i=1}^{N_{\text{conf}}} \left(\frac{\mathcal{H}_s \psi_s(\mathbf{R}_i^s)}{\psi_s(\mathbf{R}_i^s)} W_i - \frac{\mathcal{H} \psi(\mathbf{R}_i)}{\psi(\mathbf{R}_i)} \right), \quad (6)$$

where

$$W_i = \frac{N_{\text{conf}} |\psi_s(\mathbf{R}_i^s)/\psi(\mathbf{R}_i)|^2 J(\mathbf{R}_i)}{\sum_{j=1}^{N_{\text{conf}}} |\psi_s(\mathbf{R}_j^s)/\psi(\mathbf{R}_j)|^2 J(\mathbf{R}_j)}, \quad (7)$$

and $J(\mathbf{R})$ is the Jacobian for the transformation (Eq. 4).

3 Correlated sampling in diffusion Monte Carlo

In DMC [11], the primary walk is generated according to a stochastic implementation of the integral equation:

$$f(\mathbf{R}', t + \tau) = \int d\mathbf{R} G(\mathbf{R}', \mathbf{R}, \tau) f(\mathbf{R}, t), \quad (8)$$

where the importance-sampled Green's function $G(\mathbf{R}', \mathbf{R}, \tau) = \psi(\mathbf{R}') \langle \mathbf{R}' | \exp\{-\mathcal{H}\tau\} | \mathbf{R} \rangle / \psi(\mathbf{R})$, $f = \phi\psi$, ϕ is the ground state wave function and ψ the trial wave function. For small values of τ (short-time approximation), $G(\mathbf{R}', \mathbf{R}, \tau)$ is given by the product of three factors, drift, diffusion and growth/decay:

$$G(\mathbf{R}', \mathbf{R}, \tau) \approx \frac{1}{(2\pi\tau)^{\frac{3N}{2}}} e^{-\frac{(\mathbf{R}' - \mathbf{R} - \mathbf{V}(\mathbf{R})\tau)^2}{2\tau}} e^{S(\mathbf{R}', \mathbf{R}, \tau)}, \quad (9)$$

where $\mathbf{V} = \nabla\psi(\mathbf{R})/\psi(\mathbf{R})$ and $S(\mathbf{R}', \mathbf{R}, \tau) = (2E_T - E_L(\mathbf{R}') - E_L(\mathbf{R}))\tau/2$ with $E_L = \mathcal{H}\psi(\mathbf{R})/\psi(\mathbf{R})$. A set of primary walkers characterized by the pairs (\mathbf{R}_i, w_i) is a random realization of the distribution f . Each walker executes a branching random walk: a walker originally at \mathbf{R} drifts to $\mathbf{R} + \mathbf{V}(\mathbf{R})\tau$ and then diffuses to \mathbf{R}' according to the Gaussian term in Eq. 9. To ensure that, when ψ is the ground state wave function, ψ^2 is sampled exactly despite the short-time approximation in the Green's function, the move is accepted with probability

$$p = \min \left\{ 1, \frac{|\psi(\mathbf{R}')|^2 T(\mathbf{R}, \mathbf{R}', \tau)}{|\psi(\mathbf{R})|^2 T(\mathbf{R}', \mathbf{R}, \tau)} \right\}, \quad (10)$$

as prescribed by the detailed balance condition. We denote by T the drift-diffusion part of the Green's function G . Finally, the weight of the walker is multiplied by $\exp[S(\mathbf{R}', \mathbf{R}, \tau)]$. In practice, we employ the improved version of G presented in Ref. [12].

3.1 An impractical route to DMC correlated sampling

Previous attempts in the literature to perform correlated sampling in DMC adopted a simplified approach to the algorithm: they either used the same wave function for the secondary and the primary geometry, or did not include the accept/reject step, or completely neglected problems which occur at the nodes. Therefore, to clarify possible misconceptions and, at the same time, lay the building ground for our algorithm, we present in this section a correct but impractical algorithm for correlated sampling in DMC.

Let us generate the primary walk according to Eq. 9 and the secondary walk as specified by the space-warp transformation (Eq. 4). Two complications, absent in VMC, arise for correlated sampling in DMC. First of all, the dynamics of the secondary walker should have been governed by an importance sampled Green's function, $G_s(\mathbf{R}^s, \mathbf{R}^s, \tau)$, constructed from the secondary wave function ψ_s , and the move should have been accepted with probability p_s obtained by substituting, in Eq. 10, ψ and T with ψ_s and T_s , respectively.

However, the secondary geometry move was effectively proposed according to the drift-diffusion Green's function $T(\mathbf{R}', \mathbf{R}, \tau)/J(\mathbf{R}')$ and accepted with probability p defined in Eq. 10. To correct for the wrong dynamics, we should multiply the weights of the secondary walkers by

$$r \frac{G_s(\mathbf{R}^{s'}, \mathbf{R}^s, \tau)}{T(\mathbf{R}', \mathbf{R}, \tau)/J(\mathbf{R}')}, \quad (11)$$

where $r = p_s/p$ if the move is accepted and $r = (1 - p_s)/(1 - p)$ if the move is rejected. However, these products fluctuate wildly (r can be anywhere between zero and infinity). Therefore, it is not practical to follow this route to perform correlated sampling unless bounds can be placed on the ratios while at the same time ensuring that unbiased results are obtained in the $\tau \rightarrow 0$ limit.

An additional complication is the common practice in fixed-node DMC to reject moves that cross nodes. If primary and secondary walkers were to be treated on the same footing (p_s set to zero when the secondary walker crosses its own nodes), the weights of the secondary walkers would all become zero in a sufficiently long run. Even though this problem can be easily overcome since it is legitimate to do fixed-node DMC allowing walkers to cross nodes [12], reweighting as in Eq. 11 remains impractical due to the large fluctuations.

3.2 Our accurate and efficient algorithm

Our alternative correlated sampling algorithm [10] is approximate but very accurate. Given the successful implementation of correlated sampling in VMC, we want to devise a scheme which differs as little as possible from VMC and is almost as efficient as VMC. Also, to compute DMC energies for just the reference geometry, we have a well known algorithm including an accept/reject step, which is efficient and has a very small time-step error. Therefore, we wish to have a DMC algorithm which is close to our VMC algorithm for calculating energy differences, reduces to our usual DMC algorithm for the reference geometry, and yields results very close to the DMC limit for the secondary geometry.

Our algorithm is based on the observation that, in the absence of the growth/decay factor and presence of the accept/reject step, we would be sampling ψ^2 for the primary walk, and ψ_s^2 for the secondary walk by reweighting the averages with the ratio of wave functions (Eq. 7). If we stopped here, we would be simply obtaining the VMC energy difference. To approach the DMC limit, we then multiply the weights of the primary and secondary walkers by the corresponding growth/decay factors. Thus, we recover the fixed-node solution for the primary walk, but we do so only approximately for the secondary

walk since the moves were not proposed with the right dynamics. To partially correct for this, we introduce a secondary time-step as discussed below. This algorithm must of course always give energies lower than the VMC energies for the secondary geometries and we show in Sec. 5 that it is indeed very accurate. We summarize our recipe as follows:

- (1) We generate secondary walks from the reference walk according to the space-warp transformation.
- (2) In the averages for the secondary geometry, we retain the ratios of the secondary and primary wave functions as in VMC (Eqs. 6 and 7).
- (3) The secondary weights are the primary ones which have been undone and multiplied by the secondary growth/decay factor over the last N_{proj} generations:

$$w^s = w \prod^{N_{\text{proj}}} \frac{\exp[S_s(\mathbf{R}^{s'}, \mathbf{R}^s, \tau_s)]}{\exp[-S(\mathbf{R}', \mathbf{R}, \tau)]}. \quad (12)$$

The last step keeps the primary and the secondary walk correlated since the secondary and primary weights only differ over the last N_{proj} generations. N_{proj} is chosen large enough to project out the secondary ground state, but small enough to avoid a considerable increase in the fluctuations. In the exponential factors, we introduced τ_s because the secondary moves are effectively proposed with a different time-step, τ_s , in the drift-diffusion term of Eq. 9: if the molecule is for instance stretched, the secondary electronic coordinates are stretched and, to a first approximation, we are sampling a Gaussian with a larger width, or equivalently a larger time-step, in the drift-diffusion term. A sensible definition of τ_s is therefore $\tau_s = \tau \langle \Delta R_s^2 \rangle / \langle \Delta R^2 \rangle$ where ΔR is the displacement resulting from diffusion, and ΔR_s is the displacement needed to take the secondary walker from its drifted position to the position specified by the space-warp transformation. τ_s is computed over the first equilibration blocks of the DMC run.

In the limit of vanishing displacement, the difference of primary and secondary energies and its statistical error vanish linearly, so the force and its error are well behaved.

4 Secondary geometry wave functions

We considered three choices for secondary geometry wave functions:

- (1) The secondary wave functions have the same parameters $\{\mathbf{p}\}$ as the primary one but the coordinates are relative to the new nuclear positions: $\psi_s(\mathbf{R}_i, \mathbf{R}_\alpha^s) = \psi(\mathbf{R}_i, \mathbf{R}_\alpha^s, \mathbf{p}_s)$ with $\mathbf{p}_s = \mathbf{p}$, possibly with the minimal changes required to impose the cusp conditions.
- (2) The secondary geometry wave functions at warped electron positions are related to the primary ones at the original positions, $\psi_s(\mathbf{R}_i^s, \mathbf{R}_\alpha^s) = \psi(\mathbf{R}_i, \mathbf{R}_\alpha, \mathbf{p}) / \sqrt{J(\mathbf{R}_i)}$. This wave function depends on the transformation (it was used in Ref. [9, b] with a different transformation) and has the advantages that the weights W_i in (Eq. 2) are unity and that the transformation of Eq. 4 maps the nodes of the primary geometry wave function onto those of the secondary geometry wave function. Consequently, if the primary geometry walkers do not cross nodes, neither do the corresponding secondary geometry walkers. Surprisingly, it gives larger fluctuations of the energy differences than choice (1).
- (3) $\psi_s(\mathbf{R}_i, \mathbf{R}_\alpha^s) = \psi(\mathbf{R}_i, \mathbf{R}_\alpha^s, \mathbf{p}_s)$ with reoptimized parameters \mathbf{p}_s . This choice gives the smallest fluctuation of the energy differences and the best potential energy surface.

For choices (1) and (3) since the nodes of the primary wave function do not map onto the nodes of the secondary wave function, it is possible to encounter configurations for which the primary and secondary geometries (and consequently the corresponding weights) become much different from each other, resulting in a loss of correlation. This could be cured by placing bounds on the differences but in this paper we have not done this because for the systems and wave functions employed this was not a serious problem.

In reoptimizing the secondary geometry wave functions, we follow a particular procedure to keep the local energies of the primary and secondary wave functions, for each MC configuration and corresponding mapped configurations, as closely correlated as possible. The parameters in the Jastrow factor of the primary wave function are first optimized using the variance minimization method [13], i.e., the variance of the local energy is minimized over a set of MC configurations sampled from the square of the best wave function available before we start the optimization. When we reoptimize the Jastrow parameters of the secondary wave function, we employ the same configurations used to optimize the primary one, but now properly mapped through the space-warp transformation (Eq. 4). If one uses as a starting point the primary Jastrow recentered and only reoptimizes the linear parameters in the Jastrow (but not the parameter κ of Ref. [14]), only a very small number of optimization steps are needed to reoptimize the Jastrow factor of the secondary wave

function. While this scheme always works reliably for the reoptimization of the Jastrow parameters, we sometimes have difficulty keeping the primary and secondary wave functions correlated if we optimize the determinantal part as well. Therefore, for the determinantal part, we choose to simply recenter the determinants, and follow the above procedure only for the optimization of the Jastrow component. On the other hand, if the determinantal part of the wave function is obtained using a quantum chemistry package, we can simply reoptimize the determinantal part by rerunning the Hartree-Fock (HF) or multi-configuration self-consistent-field calculation (MCSCF) for the secondary geometries. However, in this case, the determinantal parts of both the primary and the secondary geometry wave functions are not optimized in the presence of the Jastrow factor.

We calculate all molecules with choice (1) but also demonstrate the superiority of choice (3) for all-electron B_2 and pseudo Si_2 . For B_2 , we reoptimize only the Jastrow part while, for Si_2 , we present also results with the determinantal part reoptimized using the quantum chemistry package GAMESS [15].

5 Results

The algorithms presented in the previous sections are tested in all-electron calculations of the first-row homonuclear dimers and in pseudopotential calculations of the Si_2 dimer.

For the all-electron dimers, the primary wave functions [14] were optimized close to the experimental bond length by the variance minimization method. The potential energy curves were obtained with correlated sampling from ten geometries, using the warp transformation and recentered secondary geometry wave functions (choice (1) above). Values of $N_{\text{proj}}\tau$ of 5-10 H^{-1} were sufficient to project out the secondary wave functions.

Before presenting results for the bond lengths and vibration frequencies, we discuss some of the tests we performed on our method. We first verified that, over a large range of atomic displacements, computing the energy difference between primary and secondary geometries with our correlated sampling method is more efficient than performing independent runs. In Fig. 1, we show the efficiency gain obtained with our VMC and DMC correlated sampling algorithm for Li_2 . The gain is given by the ratio of the root-mean-square fluctuations of the relative energy obtained with correlated sampling and the one derived from two separate runs. The gain diverges as the secondary geometry approaches the primary one, and is more than a factor of 30 even for displacements as large as ± 0.2 a.u.

To further ascertain the efficiency of our algorithm, we performed two ad-

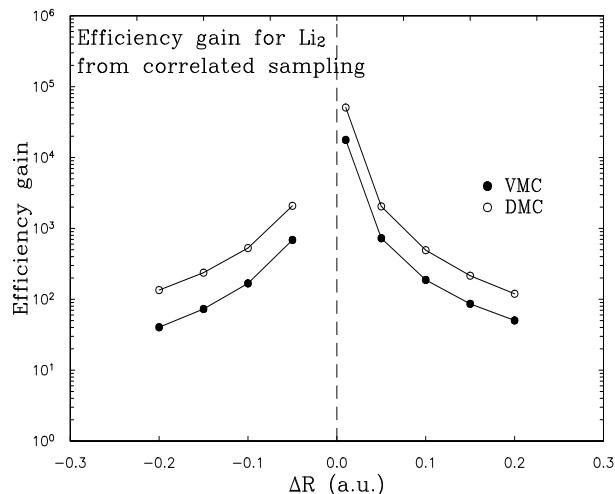


Figure 1. VMC and DMC efficiency gain for Li_2 from using correlated sampling to compute the energy difference between the primary and secondary geometries, instead of performing independent runs.

ditional calculations for B_2 ; in the first, we omitted the warp transformation, whereas in the second we employed reoptimized, rather than recentered, Jastrows in the secondary wave functions. In Fig. 2 (upper plot), we present the VMC root-mean-square fluctuations (σ_{VMC}) of the relative energy of primary and secondary geometries divided by the atomic displacement, $\Delta E/\Delta R$, for B_2 . Introducing the warp transformation yields a reduction of about a factor of 3.5-5 in σ_{VMC} , which corresponds to a factor of 12-25 saving in computer time. Moreover, σ_{VMC} is only slightly dependent on the secondary geometry used. As expected, a further reduction in σ_{VMC} is obtained when the space-warp transformation is used in combination with reoptimized, rather than recentered, secondary geometry wave functions. The space-warp transformation was found to be of even greater help for heavier molecules, e.g., for F_2 the reduction in the fluctuations was at least a factor of 3.5-10. In fact, in the absence of space-warp transformation for F_2 , it is even difficult to reliably estimate the statistical error of secondary geometries that differ considerably from the primary one.

It is clear that the gain in efficiency due to using the space-warp transformation is greater for all-electron than for pseudopotential calculations. To

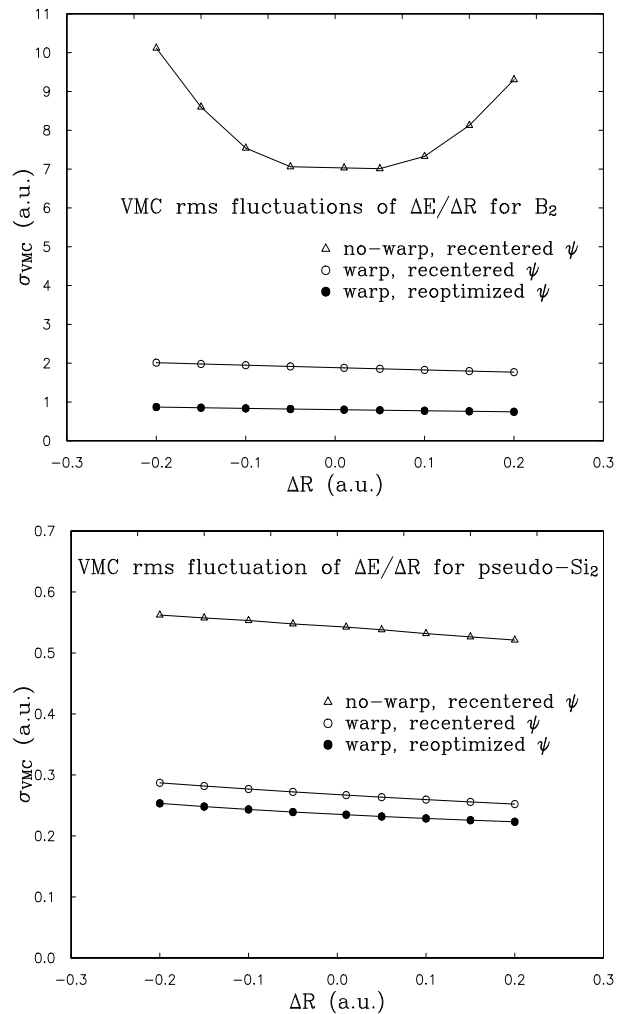


Figure 2. VMC fluctuations (σ_{VMC}) of the relative energy of the primary and secondary geometries divided by the bond stretch for all-electron B_2 (upper figure) and pseudo Si_2 (lower figure). If correlated sampling were not used, σ_{VMC} would diverge at $\Delta R = 0$. The smallest σ_{VMC} is achieved by using warping along with reoptimized secondary wave functions.

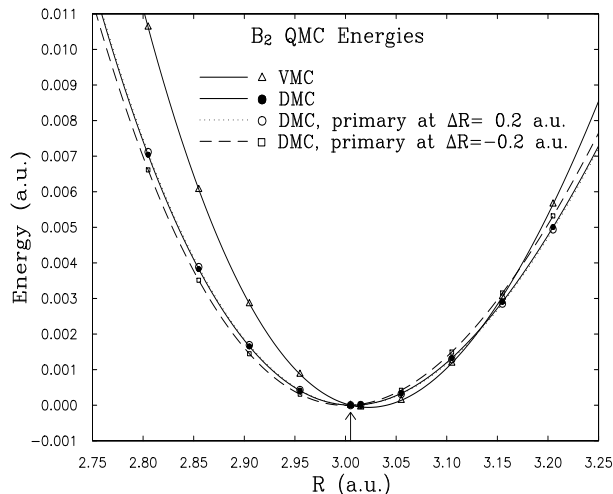


Figure 3. Potential energy curve for B_2 in VMC and DMC. The three DMC curves are obtained with three different primary geometries (equilibrium, stretched by 0.2 a.u. and -0.2 a.u.) and using recentered wave functions. All curves are shifted with the energy at the equilibrium distance (arrow) defined as the zero. Atomic units are used.

answer the question: “what, if any, gain is obtained in pseudopotential calculations”, we calculate the force $\Delta E/\Delta R$ for the Si_2 dimer and plot the corresponding σ_{VMC} in Fig. 2 (lower plot). The primary geometry wave function used is a HF determinant optimized in GAMESS, multiplied by a Jastrow factor optimized in variance minimization. The three curves are analogous to the ones for B_2 , and are obtained by recentering the wave function with and without space-warp transformation, and by reoptimizing the Jastrow factor at the secondary geometries. Employing the warp transformation yields a reduction of about a factor of 2, which corresponds to a non-negligible saving in computer time of a factor of 4. As in the all-electron case, σ_{VMC} is further reduced if the space-warp transformation is used in combination with a reoptimized secondary wave function.

To test the accuracy of our DMC correlated sampling algorithm, we performed DMC runs for H_2 and B_2 for three different primary geometries, (a) the equilibrium geometry, (b) a geometry stretched by 0.2 and (c) by -0.2 a.u, keeping the wave functions independent of the choice for the primary geometry. The runs (a), (b) and (c) should give identical potential energy

curves if the algorithm were exact. In Fig. 3, we show results for B_2 that reveal the high accuracy of our DMC algorithm: the three DMC curves are very close and clearly distinguishable from the VMC results. These results are confirmed by the calculations for H_2 where, despite the use of an intentionally poor wave function, the three curves gave the equilibrium bond lengths (a) 1.4014(2) (b) 1.4014(2) and (c) 1.4015(2) a.u. The true equilibrium bond length, from a careful fit to the results of Ref. [16], is 1.4011 a.u.

To test the improvement resulting from employing $\tau_s \neq \tau$, we performed, for H_2 , DMC correlated sampling with $\tau_s = \tau$. Since $\tau_s > \tau$ for $\Delta R > 0$ and $\tau_s < \tau$ for $\Delta R < 0$, we expect this potential energy curve to yield an equilibrium bond-length that is too short. The equilibrium bond length is indeed 1.4003(2) a.u., which is 4 standard deviations from the true bond-length, whereas that obtained with our $\tau_s \neq \tau$ algorithm, 1.4014(2) a.u., is 1.5 standard deviation from the true bond-length.

Having ascertained the accuracy and efficiency of our algorithm, we computed the bond lengths of all first-row dimers with VMC and DMC correlated sampling. In Table 1, we list the errors in the bond lengths obtained from restricted Hartree-Fock (RHF) [17], LDA [18], GGA [19], VMC and DMC. The RHF results show the worst agreement with experiment, with Be_2 not being bound. The DMC errors are, in all cases, either smaller than or comparable to those from VMC, and the root-mean-square deviations from experiments in DMC are smaller than in LDA and GGA by a factor of 3.9 and 2.6, respectively.

Table 1. Experimental bond lengths [20] of first-row dimers and theoretical errors in RHF [17], LDA [18], GGA [19], VMC and DMC (in a.u.).

molecule	Expt.	RHF	LDA	GGA	VMC	DMC
Li_2	5.051	0.270	0.069	0.057	0.101(2)	0.018(3)
Be_2	4.630	–	-0.109	-0.001	-0.069(3)	-0.014(5)
B_2	3.005	0.086	0.025	0.042	0.018(2)	0.002(2)
C_2	2.348	-0.007	0.006	0.023	0.006(2)	0.008(1)
N_2	2.074	-0.061	-0.006	0.011	0.012(2)	0.007(1)
O_2	2.282	-0.107	-0.012	0.044	0.028(2)	0.023(4)
F_2	2.668	-0.161	-0.053	0.040	0.021(4)	0.015(5)
rms	–	∞	0.054	0.036	0.049	0.014

If one is interested not only in the force but also in the vibration fre-

Table 2. Error in the bond lengths, Δr_e (in a.u.), and vibration frequencies, Δf (in cm^{-1}) of Si_2 from VMC and DMC using various trial wave functions. We indicate whether the HF wave function or a MCSCF wave function was used for the antisymmetric determinantal part, and whether the determinantal (D) and the Jastrow (J) parts of the secondary geometry wave functions were recentered or reoptimized. The numbers in parentheses are the statistical errors. The experimental bond length is 4.244 a.u. and the experimental frequency is 510.94 cm^{-1} [20].

	Wave function		Δr_e		Δf	
	D	J	VMC	DMC	VMC	DMC
HF	recent.	recent.	-0.039 (5)	-0.017 (5)	109 (7)	48 (7)
HF	reopt.	recent.	0.000 (5)	0.006 (5)	33 (7)	11 (7)
HF	reopt.	reopt.	0.020 (5)	0.007 (5)	40 (7)	3 (7)
MCSCF	recent.	recent.	0.005 (5)	0.004 (5)	78 (7)	31 (7)
MCSCF	reopt.	recent.	0.033 (5)	0.013 (5)	-5 (7)	-7 (7)
MCSCF	reopt.	reopt.	0.024 (5)	0.013 (5)	14 (7)	1 (7)

quency, then it is necessary to reoptimize at least the determinantal part of the secondary geometry wave functions for an accurate estimate of the second derivative of the energy with respect to atomic displacement. We demonstrate this for the Si_2 molecule in Table 2, where we give the VMC and DMC equilibrium bond lengths and the vibration frequencies for six wave functions. The wave functions employ either a HF or a MCSCF determinantal part. (For the MCSCF we included the 45 most important determinants resulting from single and double excitations within the s and p valence atomic shells.) The primary geometry was chosen to be the experimental equilibrium bond length and the Jastrow factor of the primary geometry wave functions was optimized by variance minimization. The VMC energies at the primary geometry are $-7.637(1)$ Hartree and $-7.633(1)$ Hartree for the HF- and the MCSCF-Jastrow wave function, respectively. It may seem surprising that the MCSCF-Jastrow wave function has a slightly higher energy than the HF-Jastrow one. The reason is that the determinantal part of the wave function was not reoptimized in the presence of the Jastrow.

One expects that, if the secondary geometry wave functions are not reoptimized, the calculated vibration frequencies will be too high. The numbers in Table 2 confirm this expectation but they also indicate that it is probably sufficient to reoptimize just the determinantal part of the secondary geometry wave functions. There is a clear tendency for bond lengths from calculations

employing the MCSCF determinants to be longer than those from calculations employing the HF determinant. The reason for this is that the MCSCF determinants are better able to describe the dissociated molecule than the HF determinant and consequently the energies at long bond lengths are lowered relative to the energies at short bond lengths.

6 Comparison to variance-reduced Hellman-Feynman method

Unfortunately, there has not been any study of the relative efficiency of our correlated sampling technique with the variance-reduced correlated sampling technique of Assaraf and Caffarel [6] but some statements can be made about the merits of each method.

The Hellman-Feynman expression for the q^{th} component of the force is

$$\langle F_q \rangle = \int d\mathbf{R} \psi(\mathbf{R}) F_q(\mathbf{R}) \psi(\mathbf{R}) / \int d\mathbf{R} |\psi(\mathbf{R})|^2, \quad (13)$$

where $F_q(\mathbf{R}) = -\nabla_q V(\mathbf{R})$ and $V(\mathbf{R})$ is the potential energy. This expression, used in VMC, has an infinite variance: $\langle F_q(\mathbf{R})^2 \rangle$ is infinite since $F_q(\mathbf{R})$ diverges as $1/r^2$, where r is the distance of an electron from the nucleus. Instead, Assaraf and Caffarel calculate the expectation value of

$$\tilde{F}_q(\mathbf{R}) = F_q(\mathbf{R}) + \left[\frac{\tilde{H}\tilde{\psi}}{\tilde{\psi}} - \frac{\tilde{H}\psi}{\psi} \right] \frac{\tilde{\psi}}{\psi}, \quad (14)$$

where \tilde{H} is Hermitian and consequently the expectation value of the additional terms is zero for Monte Carlo configurations sampled from $|\psi|^2$. They make the simple choice $\tilde{H} = H$ and $\tilde{\psi} = Q\psi$, where Q is chosen such that the leading divergence of $F_q(\mathbf{R})$ is cancelled. Consequently, $\tilde{F}_q(\mathbf{R})$ diverges only as $1/r$ and $\langle \tilde{F}_q(\mathbf{R})^2 \rangle$ is finite. The remaining $1/r$ divergence in $F_q(\mathbf{R})$ could be possibly removed with an appropriate choice of $\tilde{\psi}$ [21].

The advantage of this modified Hellman-Feynman method is that it is not necessary to calculate reoptimized secondary geometry wave functions and the Hamiltonian acting on them. Consequently, it becomes more efficient than the correlated sampling method if many components of the force need to be calculated. The disadvantage of this Hellman-Feynman method is that it is not a zero-variance method in the limit that the trial wave function ψ becomes the true wave function. Another disadvantage is that the fluctuations of the force grow rapidly with increasing nuclear charge because Q is proportional to the nuclear charge Z and also because, for large Z , electrons are on average closer to the nucleus. The statistical errors of the correlated sampling technique also grow with Z but not as rapidly. Unfortunately, we cannot make

this comparison more quantitative since insufficient details of the lengths of the Monte Carlo runs and the statistical errors are provided in Ref. [6]. Yet another disadvantage is that since the distribution sampled in DMC is the mixed distribution, $\psi_{\text{FN}}\psi$, and since the Hellman-Feynman operator does not commute with the Hamiltonian, the force calculated from the usual “mixed estimator” has an error that is linear in the difference between ψ_{FN} and ψ . A force with an error that is quadratic in the difference was calculated by Assaraf and Caffarel from the extrapolated estimator $2\mathbf{F}_{\text{DMC}} - \mathbf{F}_{\text{VMC}}$ but this leads to a somewhat larger statistical error. Alternatively, the forward walking technique [2] can be used to eliminate the error but this increases the statistical error even more.

7 Conclusions

We presented an efficient method to compute numerical forces and vibration frequencies in DMC, a long-standing unsolved problem in QMC techniques. The method is very accurate and was tested on first-row dimers, where the DMC bond lengths were found to agree with experiment better than those from HF, LDA, GGA and VMC.

Acknowledgments

This work began during a visit to the Institute for Nuclear Theory in Seattle and is funded by Sandia National Laboratory. Sandia is a multiprogram laboratory operated by Sandia Corporation, a Lockheed Martin Company, for the United States Department of Energy, under Contract No. DEAC0494AL85000.

References

1. D. M. Ceperley and B. J. Alder, Phys. Rev. B **36**, 2092 (1987); S. Fahy, X. W. Wang, and S. G. Louie, Phys. Rev. B **42**, 3503 (1990); J. C. Grossman and L. Mitas, Phys. Rev. Lett. **79**, 4353 (1997); R. Q. Hood, M. Y. Chou A. J. Williamson, G. Rajagopal, and R. J. Needs, Phys. Rev. B **57**, 8972 (1998); Arne Lüchow and J. B. Anderson, Annu. Rev. Phys. Chem., **51**, 501 (2000).
2. W. M. C. Foulkes, L. Mitas, R. J. Needs, and G. Rajagopal, Rev. Mod. Phys. **73**, 33 (2001); B. L. Hammond, W. A. Lester, and P. J. Reynolds, *Monte Carlo Methods in an initio quantum chemistry* (World Scientific,

- 1994); M. H. Kalos and P. A. Whitlock, *Monte Carlo Methods, Vol. 1* (Wiley-Interscience, 1986).
3. H. Hellman, *Einführung in die Quanten Theorie* (Deuticke, Leipzig, 1937); R. P. Feynman, Phys. Rev. **56**, 340 (1939).
 4. P. Pulay, Mol. Phys. **17**, 197 (1969); M. Scheffler, J. P. Vigneron and G. B. Bachelet, Phys. Rev. B **31**, 6541 (1985).
 5. K. C. Huang, R. J. Needs, and G. Rajagopal, J. Chem. Phys. **112**, 4419 (2000).
 6. R. Assaraf and M. Caffarel, Phys. Rev. Lett. **83**, 4682 (1999); J. Chem. Phys. **113**, 4028 (2000).
 7. R. E. Lowther and R. L. Coldwell, Phys. Rev. A **22**, 14 (1980).
 8. C. J. Umrigar, Int. J. Quant. Chem. Symp. **23**, 217 (1989).
 9. a) P. J. Reynolds, R. N. Barnett, B. L. Hammond, R. M. Grimes, and W. A. Lester, Int. J. Quant. Chem. **29**, 589 (1986); b) C. A. Traynor and J. B. Anderson, Chem. Phys. Lett. **147**, 389 (1988); c) J. Vrbik and S. M. Rothstein, J. Chem. Phys. **96**, 2071 (1992).
 10. C. Filippi and C. J. Umrigar, Phys. Rev. B **61**, R16291 (2000).
 11. J. B. Anderson, J. Chem. Phys. **63**, 1499 (1975); *ibid.* **65**, 4121 (1976); P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester, *ibid.* **77**, 5593 (1982).
 12. C. J. Umrigar, M. P. Nightingale, and K. J. Runge, J. Chem. Phys. **99**, 2865 (1993).
 13. C. J. Umrigar, K. G. Wilson, and J. W. Wilkins, Phys. Rev. Lett. **60**, 1719 (1988).
 14. C. Filippi and C. J. Umrigar, J. Chem. Phys. **105**, 213 (1996); for O₂, we used an improved wave function.
 15. M. W. Schmidt, K. K. Baldrige, J. A. Boatz, S. T. Albert, M. S. Gordon, J. H. Jensen, S. Koseki, N. Matsunaga, K. A. Nguyen, S. Su, T. L. Windus, M. Dupuis, and J. A. Montgomery, Jr., J. Comput. Chem. **14**, 1347 (1993).
 16. L. Wolniewicz, J. Chem. Phys. **99**, 1851 (1993).
 17. K. A. Peterson, R. A. Kendall, and T. H. Dunning, J. Chem. Phys. **99**, 9790 (1993).
 18. A. D. Becke, Phys. Rev. A **33**, 2786 (1986); R. M. Dickson and A. D. Becke, J. Chem. Phys. **99**, 3898 (1993).
 19. F. W. Kutzler and G. S. Painter, Phys. Rev. B **45**, 3236 (1992); they employ the Perdew-Wang '86 GGA.
 20. K. P. Huber and G. Herzberg, *Molecular Spectra and Molecular Structure. IV. Constants of Diatomic Molecules* (Van Nostrand, Princeton, 1979); for Be₂, V. E. Bondybey and J. H. English, J. Chem. Phys. **80**, 568

(1984).

21. M. Caffarel, private communication.

Index

bond lengths of first-row dimers, 13

correlated sampling

in DMC, 4–7

in VMC, 3

optimization of wave function, 8

space-warp coordinate transformation, 4

diffusion Monte Carlo, 1–16

force

DMC and VMC bond lengths of first-row dimers, 13

variance of, 9

frequency, 13

generalized gradient approximation, 3

Hartree-Fock, 3

Hellman-Feynman theorem, 1–2, 15–16

applicability of, 1–2

reduced variance method, 15–16

local density approximation, 3