PHYSICAL REVIEW B **86**, 201107(R) (2012)

Efficient calculation of unbiased expectation values in diffusion quantum Monte Carlo

Manolo C. Per,[1,2,*] Ian K. Snook,[2] and Salvy P. Russo[2]

[1]*Virtual Nanoscience Laboratory, CSIRO Materials Science and Engineering, Parkville, Victoria 3052, Australia*

[2]*Applied Physics, School of Applied Sciences, RMIT University, Melbourne 3001, Australia*

(Received 4 May 2012; revised manuscript received 29 June 2012; published 29 November 2012)

Despite the proven utility of quantum Monte Carlo methods in addressing the quantum many-body problem, many important observables are difficult to calculate due to the presence of large, and sometimes divergent, statistical errors. The present state of the art allows the construction of renormalized estimators which result in finite variances, but which invariably include some systematic bias. We present a simple method for calculating unbiased expectation values of local operators in the diffusion quantum Monte Carlo method which is applicable to both bare and renormalized estimators, allowing the accurate calculation of important properties such as forces.

DOI: 10.1103/PhysRevB.86.201107 PACS number(s): 05.10.Ln, 02.70.Ss, 03.65.−w

The diffusion quantum Monte Carlo (DMC) method is a powerful approach to solving the quantum many-body problem. It can be used to calculate numerically exact ground-state energies for bosonic systems, and highly accurate fermionic energies using the fixed-node approximation.[1] Calculating accurate expectation values for general observables is more difficult than for total energies due to the presence of systematic bias and large statistical variance. Indeed, a number of important estimators have formally infinite variance, the most well-known example being forces.[2,3] Individually, solutions to these two problems of bias and variance exist. Forward-walking[4] and reptation Monte Carlo[5] (although strictly a different approach to DMC) can be used to remove the systematic bias, while the renormalized estimator approach introduced by Assaraf and Caffarel[6,7] can be used to reduce the statistical variance. Unfortunately, these approaches are incompatible. The forward-walking and reptation approaches work by sampling the pure distribution Ψ_0^2 rather than the usual DMC mixed distribution $\Psi_0\Psi_T$, where Ψ_T is a user-defined trial wave function. In contrast, calculating unbiased expectation values of renormalized estimators requires a very different quantity, namely, the derivative of the mixed distribution $\partial_\lambda(\Psi_0\Psi_T)$. As a result, calculations using renormalized estimators always contain some systematic bias.

In this Rapid Communication we present a simple method which enables the calculation of unbiased expectation values of renormalized estimators, and is also applicable to bare (i.e., nonrenormalized) estimators. Our approach is based on the fact that expectation values can be written as derivatives of the total energy. The expectation value of a bare local operator $\hat{O}$ can be evaluated by perturbing the Hamiltonian $\hat{H} \to \hat{H} + \lambda\hat{O}$, and calculating the derivative of the total energy with respect to λ at $\lambda = 0$. Applying this approach to the total DMC energy we can write the unbiased expectation value as

$$\langle \hat{O} \rangle_{\Psi_0^2} = \partial_\lambda E_{\mathrm{DMC}} = \frac{\int \Psi_0 \Psi_T \partial_\lambda E_L d\mathbf{r}}{\int \Psi_0 \Psi_T d\mathbf{r}} + \frac{\int \partial_\lambda (\Psi_0\Psi_T)(E_L - E_{\mathrm{DMC}})\, d\mathbf{r}}{\int \Psi_0 \Psi_T d\mathbf{r}}, \quad (1)$$

where the local energy is $E_L = \hat{H}\Psi_T/\Psi_T$. This expression can be used to define low variance "renormalized" estimators through a suitable choice of the trial first-order wave function $\partial_\lambda \Psi_T$.[2,8,9] The problem with using Eq. (1) is that the exact first-order wave function $\partial_\lambda \Psi_0$, which appears in the second integral, is not available in standard DMC calculations. In practice the approximation $\Psi_T \partial_\lambda \Psi_0 \approx \Psi_0 \partial_\lambda \Psi_T$ is often used, resulting in a systematic bias. Here we show that the second integral in Eq. (1), referred to as the zero-bias (ZB) integral, can be evaluated exactly (within the fixed-node approximation) by extending the DMC method itself to first order in perturbation theory to obtain the derivative of the mixed distribution directly. This results in a true zero-bias DMC (ZBDMC) method, in which expectation values are independent of Ψ_T.

We begin by reviewing some relevant aspects of DMC. The DMC method can be written as an integral evolution equation,

$$f(\mathbf{r}', t+\tau) = \int \tilde{G}(\mathbf{r}, \mathbf{r}'; \tau) f(\mathbf{r}, t) d\mathbf{r}, \quad (2)$$

which describes the evolution of an initial probability distribution $f(\mathbf{r}, t)$ from time t to time $t + \tau$ under the action of an importance-sampled Green's function $\tilde{G}$. In the long-time limit, as $\tau \to \infty$, the distribution evolves to become proportional to the mixed distribution,

$$\lim_{\tau\to\infty} f(\mathbf{r}', t+\tau) = \lim_{\tau\to\infty} \Psi_T(\mathbf{r}')\Psi_0(\mathbf{r}') e^{-\tau(E_0 - E_T)} \langle \Psi_0 | \phi \rangle, \quad (3)$$

where $\phi(\mathbf{r}) = f(\mathbf{r}, t)/\Psi_T(\mathbf{r})$ and E_T is a trial energy used to stabilize the calculation.

A suitable approximation to the Green's function, which is exact in the limit $\tau \to 0$, is given by the short-time approximation

$$\tilde{G}_{\mathrm{st}}(\mathbf{r}, \mathbf{r}'; \tau) = C \exp\left(-\frac{[\mathbf{r}' - \mathbf{r} - \tau \mathbf{v}(\mathbf{r})]^2}{2\tau}\right) \times \exp\left(-\frac{\tau}{2}[E_L(\mathbf{r}') + E_L(\mathbf{r}) - 2E_T]\right), \quad (4)$$

where C is a normalization constant, and $\mathbf{v} = \nabla\Psi_T/\Psi_T$. The first exponential is a stochastic kernel which is interpreted as a drift-diffusion term. The second exponential is a non-negative weight, usually referred to as the branching term.

In order to reach the long-time limit, the short-time Green's function $\tilde{G}_{\mathrm{st}}$ is repeatedly applied using a small value of τ. For

M steps the evolution is written

$$f(\mathbf{r}_M,t+M\tau)=\int \tilde{G}_{\rm st}(\mathbf{r}_1,\mathbf{r}_2;\tau)\cdots\tilde{G}_{\rm st}(\mathbf{r}_{M-1},\mathbf{r}_M;\tau) \times f(\mathbf{r}_1,t)d\mathbf{r}_1\cdots d\mathbf{r}_{M-1} =\int\prod_{i=1}^{M-1}\tilde{G}_{\rm st}(\mathbf{r}_i,\mathbf{r}_{i+1};\tau)f(\mathbf{r}_1,t)d\mathbf{r}_1\cdots d\mathbf{r}_{M-1}. \tag{5}$$

The distribution f, which is asymptotically proportional to the mixed distribution $\Psi_0\Psi_T$, is represented by an ensemble of weighted random walkers whose motion is directed by the drift-diffusion kernel. As a result of the repeated application of the short-time Green's function, the weight associated with a walker labeled i at time step M is the product of the weights from all previous steps,

$$W_{Mi}^{(0)}=\prod_{j=1}^{M-1}\exp\left(-\frac{\tau}{2}[E_L(\mathbf{r}_{j+1})+E_L(\mathbf{r}_j)-2E_T]\right). \tag{6}$$

The DMC expectation value of the bare estimator $\hat{O}$ over the mixed distribution is calculated as the weighted sum

$$\langle\hat{O}\rangle_{\rm DMC}=\frac{\sum_i^{N_s}\sum_j^{N_w}W_{ij}^{(0)}\hat{O}_{ij}}{\sum_i^{N_s}\sum_j^{N_w}W_{ij}^{(0)}}, \tag{7}$$

where the sums are over all N_w walkers which exist at each of the N_s time steps.

Now we can move on to show how the complete derivative of the energy in Eq. (1) can be calculated. The second integral requires the derivative of the DMC mixed distribution. We can obtain this quantity by differentiating the DMC evolution equation,

$$\partial_\lambda f(\mathbf{r}',t+\tau)=\int\partial_\lambda\tilde{G}(\mathbf{r},\mathbf{r}';\tau)f(\mathbf{r},t)d\mathbf{r} +\int\tilde{G}(\mathbf{r},\mathbf{r}';\tau)\partial_\lambda f(\mathbf{r},t)d\mathbf{r}. \tag{8}$$

In the long-time limit this becomes

$$\lim_{\tau\to\infty}\partial_\lambda f=\lim_{\tau\to\infty}\partial_\lambda[\Psi_T(\mathbf{r}')\Psi_0(\mathbf{r}')]e^{-\tau(E_0-E_T)}\langle\Psi_0|\phi\rangle +\lim_{\tau\to\infty}\Psi_T(\mathbf{r}')\Psi_0(\mathbf{r}')e^{-\tau(E_0-E_T)} \times[\partial_\lambda\langle\Psi_0|\phi\rangle-\tau(\partial_\lambda E_0-\partial_\lambda E_T)\langle\Psi_0|\phi\rangle]. \tag{9}$$

The first term in this limit is proportional to the derivative of the DMC mixed distribution, which is the quantity that we require. There are also contamination terms, which are proportional to the DMC mixed distribution and therefore do not contribute to the value of the ZB integral. In order to understand where the contamination comes from, we note that while the DMC method is effectively solving the Schrödinger eigenvalue equation $(\hat{H}-E)\Psi_0=0$, the first-order DMC method is effectively solving the Sternheimer[10] equation

$$(\hat{H}-E)\partial_\lambda\Psi_0=-(\partial_\lambda\hat{H}-\partial_\lambda E)\Psi_0. \tag{10}$$

This equation is weakly singular, in that if $\partial_\lambda\Psi_0$ is a solution, then so is $\partial_\lambda\Psi_0+c\Psi_0$ for any constant c.

The first-order evolution equation in Eq. (8) requires the derivative of the Green's function. The derivative of the short-time Green's function is

$$\partial_\lambda\tilde{G}_{\rm st}(\mathbf{r},\mathbf{r}';\tau)=\tilde{G}_{\rm st}(\mathbf{r},\mathbf{r}';\tau)\Big\{[\mathbf{r}'-\mathbf{r}-\tau\mathbf{v}(\mathbf{r})]\cdot\partial_\lambda\mathbf{v}(\mathbf{r}) -\frac{\tau}{2}[\partial_\lambda E_L(\mathbf{r}')+\partial_\lambda E_L(\mathbf{r})-2\partial_\lambda E_T]\Big\}. \tag{11}$$

This expression can be interpreted in a similar way to the short-time Green's function itself. The stochastic drift-diffusion kernel and branching weight from the original Green's function are retained, and we now have an additional weight, given by the term in braces, which can be either positive or negative. The first term in this new weight is the contribution from the drift-diffusion part of the Green's function, while the second term is the contribution from the original branching weight.

Despite the similarity of the interpretation of the derivative of the Green's function to the original Green's function, the new weights behave quite differently from the original branching weights. The derivative of the evolution equation using the repeated application of the short-time Green's function can be written in a natural way as

$$\partial_\lambda f(\mathbf{r}_M,t+M\tau) =\int\prod_{i=1}^{M-1}\tilde{G}_{\rm st}(\mathbf{r}_i,\mathbf{r}_{i+1};\tau)\sum_{j=1}^{M-1}\frac{\partial_\lambda\tilde{G}_{\rm st}(\mathbf{r}_j,\mathbf{r}_{j+1};\tau)}{\tilde{G}_{\rm st}(\mathbf{r}_j,\mathbf{r}_{j+1};\tau)} \times f(\mathbf{r}_1,t)d\mathbf{r}_1\cdots d\mathbf{r}_{M-1}. \tag{12}$$

We have neglected the term involving the derivative of the initial distribution, as this can (without loss of generality) be assumed to be independent of the perturbation. Our first-order distribution $\partial_\lambda f$ can be represented in the same way as the distribution f, as a collection of weighted delta functions. Equation (12) shows that the new weights are additive, rather than being multiplicative as the original branching weights. Thus, the weight associated with a walker labeled i at time step M is now the product of the original weight [shown in Eq. (6)] and a new weight

$$W_{Mi}^{(1)}=\sum_{j=1}^{M-1}\Big\{[\mathbf{r}_{j+1}-\mathbf{r}_j-\tau\mathbf{v}(\mathbf{r}_j)]\cdot\partial_\lambda\mathbf{v}(\mathbf{r}_j) -\frac{\tau}{2}[\partial_\lambda E_L(\mathbf{r}_{j+1})+\partial_\lambda E_L(\mathbf{r}_j)-2\partial_\lambda E_T]\Big\}, \tag{13}$$

where the sum is over all previous time steps. Using the new weight, we can obtain an unbiased estimate of the ZB integral in Eq. (1) using

$$\frac{\sum_i^{N_s}\sum_j^{N_w}W_{ij}^{(0)}W_{ij}^{(1)}(E_{L\,ij}-E_{\rm DMC})}{\sum_i^{N_s}\sum_j^{N_w}W_{ij}^{(0)}}. \tag{14}$$

So far we have not considered the effect of using the fixed-node approximation (FNA), which is necessary in practical calculations involving many electrons. The simplest way of implementing the FNA is to delete walkers which attempt to cross the nodal surface. This can be described in the context of the evolution equation by replacing the short-time Green's

function describing a single step with the product

$$\tilde{G}_{\rm st}^{\rm FNA}(\mathbf{r},\mathbf{r}';\tau) = \tilde{G}_{\rm st}(\mathbf{r},\mathbf{r}';\tau)\alpha(\mathbf{r},\mathbf{r}'). \quad (15)$$

The $\alpha(\mathbf{r},\mathbf{r}')$ term is a weight which is equal to 0 if the move from $\mathbf{r}$ to $\mathbf{r}'$ crosses a node, and 1 otherwise. The derivative of the FNA Green's function is

$$\partial_\lambda \tilde{G}_{\rm st}^{\rm FNA}(\mathbf{r},\mathbf{r}';\tau) = \partial_\lambda \tilde{G}_{\rm st}(\mathbf{r},\mathbf{r}';\tau)\alpha(\mathbf{r},\mathbf{r}'). \quad (16)$$

Note that there is no contribution from the derivative of the fixed-node weight $\partial_\lambda\alpha$. The reason for this is that in order for the derivative of the DMC energy to be equal to the (fixed-node) unbiased expectation value of the operator $\hat{O}$, it is necessary that the perturbation does not affect the nodal surface.[11] As a result, the new weight defined in Eq. (13) is unchanged. A more common approach to implementing the fixed-node approximation is to reject moves which attempt to cross the nodal surface. In this case, for rejected moves there is no drift-diffusion contribution to the weight $W^{(1)}$.

We have applied our ZBDMC method to calculate forces in diatomic molecules as a function of bond length. Tests on the H_2 molecule show the method is able to reproduce the exact forces, even when using an extremely poor trial wave function consisting of a single $1s$-type function on each nucleus. Results from force calculations on the BH molecule using a low-quality trial wave function are shown in Fig. 1. As the bare force estimator has infinite variance, we use the renormalized estimator defined in Ref. 2. Forces obtained with this estimator are not strictly exact, as they neglect the change in the nodal surface caused by perturbing atomic positions. In the present case this effect turns out to be tiny, but we stress that it is possible to construct a force estimator which includes these nodal effects by choosing $\partial_\lambda\Psi_T$ to be the derivative of a properly cusp-corrected trial wave function.[3] In addition to the true ZBDMC result, we present results calculated using the mixed distribution (defined by setting $\partial_\lambda\Psi_0 = 0$) and the usual approximate zero-bias method (defined by setting $\Psi_T\partial_\lambda\Psi_0 = \Psi_0\partial_\lambda\Psi_T$). The forces calculated with the approximate methods contain a significant amount of systematic bias due to the poor trial wave function used, whereas the ZBDMC results accurately reproduce the

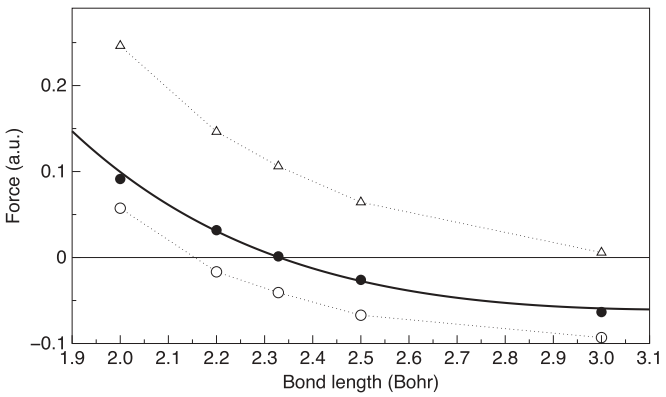

FIG. 1. Force on the H atom in a BH molecule as a function of bond length. Open triangles: DMC mixed estimate. Open circles: Approximate ZB estimate. Solid circles: True ZBDMC value. Statistical errors are smaller than the symbols. The solid black curve was obtained by differentiating a Morse potential fit to accurate DMC energies.

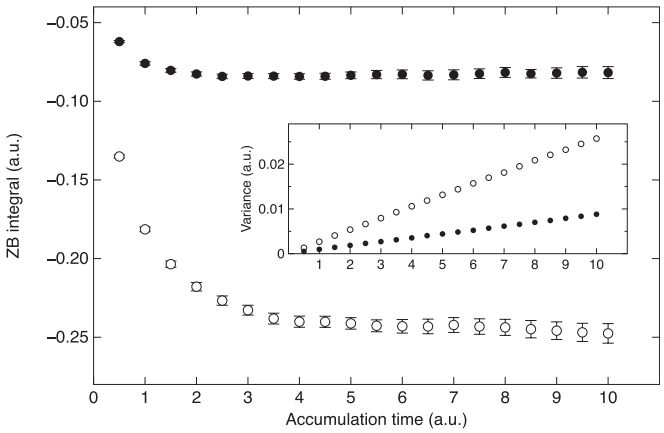

FIG. 2. Behavior of the ZB term and its variance (see inset) as a function of accumulation time, for trial wave functions of different quality in the force calculation of BH. Open circles indicate a poorer trial wave function, and solid circles a better wave function.

forces obtained by differentiating a Morse potential fit to DMC energies. The ZBDMC equilibrium bond length agrees closely with the experimental value of 2.3285 bohrs.

Some important aspects of the behavior of the estimate of the ZB integral in Eq. (1) are shown in Fig. 2. The variance of the ZBDMC estimate increases linearly with the length of the sum used to determine the new weights $W^{(1)}$. To deal with this, we simply restrict the back sum in Eq. (13) to be over a fixed finite number of steps, referred to as the accumulation time. This leads to a small amount of bias in the result, which decays with increasing accumulation time. A side effect of using a finite accumulation time is that the derivative of the trial energy in Eq. (13) (which can be chosen to be a constant) does not contribute to the value of the ZB integral, so in all our calculations we use $\partial_\lambda E_T = 0$. The use of a better trial wave function leads to a quicker convergence to the unbiased result, and a shallower gradient in the behavior of the variance. The absolute value of the ZB integral is smaller when a more accurate trial wave function is used, and it vanishes entirely when the trial wave function is exact.

Although the primary motivation for using our ZBDMC method is to calculate unbiased expectation values of renormalized estimators, it can also be used with standard bare estimators by choosing $\partial_\lambda\Psi_T = 0$. In that case, the contribution to the new weight $W^{(1)}$ from the derivative of the drift-diffusion term vanishes, and we are left with a contribution from the derivative of the branching weight, which involves the bare estimator,

$$W_{Mi}^{(1)} = -\frac{\tau}{2}\sum_{j=1}^{M-1}[\hat{O}(\mathbf{r}_{j+1}) + \hat{O}(\mathbf{r}_j)]. \quad (17)$$

When this weight is used in Eq. (14) to evaluate the ZB integral in Eq. (1), we obtain an expression similar to that of Gaudoin and Pitarke.[13] Thus, our ZBDMC method contains their Hellmann-Feynman sampling (HFS) method as a special case, which is valid only when using nonrenormalized estimators.

While the ZBDMC method can be used with bare estimators, there are definite advantages to using renormalized estimators even when the bare estimators have finite variance.

To illustrate this point, we have calculated the electron-nucleus Coulomb potential (V_{eN}) in the Be atom using both bare and renormalized estimators. The bare estimator is the standard expression, $-Z\sum_i r_i^{-1}$, and the renormalized estimator is given by

$$\partial_\lambda \Psi_T = -Z\Psi_T \sum_i r_i, \tag{18}$$

where Z is the nuclear charge and r_i is the distance between electron i and the nucleus.

Figure 3 shows the ZBDMC results using both the bare and renormalized estimators as a function of the accumulation time used to calculate the $W^{(1)}$ weights. The renormalized estimator shows a linear increase in variance as the accumulation time is increased, in the same manner as shown for the force in Fig. 2. In contrast, using the bare estimator results in a much more rapid nonlinearly increasing variance for long accumulation times. The time-step error associated with the bare estimator is much larger than that of the renormalized estimator, though both estimators do converge to the same value in the $\tau \to 0$ limit. Using a more complex DMC algorithm such as the one described in Ref. 14 should reduce the time-step errors of both estimates. The extension of the ZBDMC method to this type of algorithm will be presented in a future paper.

To conclude, we have presented a method which allows the calculation of unbiased expectation values in DMC. The method is inherently linked to the use of renormalized estimators, and so allows the efficient and accurate calculation of quantities such as forces. We have shown that the Hellmann-Feynman-sampling method of Ref. 13 corresponds to applying our method to bare rather than renormalized estimators. We have also shown that attempting to calculate the unbiased expectation value of bare estimators (even those with finite variance) leads to a rapidly increasing variance. In contrast, when using renormalized estimators, the variance increases only linearly with the accumulation time.

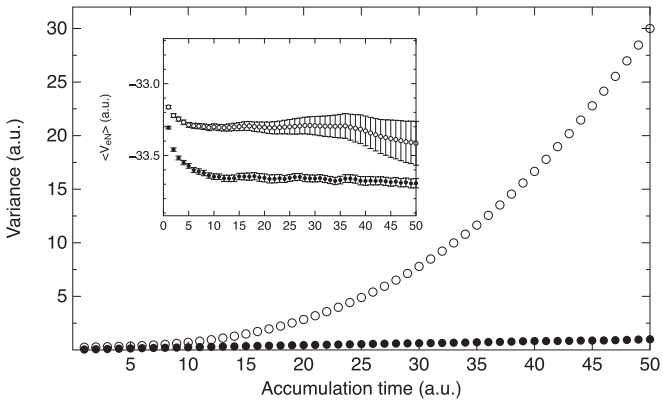

FIG. 3. Comparison of the behavior of the ZBDMC variance and expectation value (see inset) of V_{eN} in the Be atom using the bare estimator (open circles) and the renormalized estimator (solid circles). All calculations used a time step of 0.001 a.u. and a target population of 2048 walkers. The exact expectation value is close to −33.711, from Ref. 12.

*manolo.per@csiro.au

[1]J. B. Anderson, J. Chem. Phys. **65**, 4121 (1976).
[2]R. Assaraf and M. Caffarel, J. Chem. Phys. **113**, 4028 (2000).
[3]M. C. Per, S. P. Russo, and I. K. Snook, J. Chem. Phys. **128**, 114106 (2008).
[4]K. Liu, M. Kalos, and G. Chester, Phys. Rev. A **10**, 303 (1974).
[5]S. Baroni and S. Moroni, Phys. Rev. Lett. **82**, 4745 (1999).
[6]R. Assaraf and M. Caffarel, Phys. Rev. Lett. **83**, 4682 (1999).
[7]R. Assaraf and M. Caffarel, J. Chem. Phys. **119**, 10536 (2003).
[8]J. Toulouse, R. Assaraf, and C. J. Umrigar, J. Chem. Phys. **126**, 244112 (2007).
[9]M. C. Per, I. K. Snook, and S. P. Russo, J. Chem. Phys. **135**, 134112 (2011).
[10]R. Sternheimer, Phys. Rev. **84**, 244 (1951).
[11]K. Huang, R. J. Needs, and G. Rajagopal, J. Chem. Phys. **112**, 4419 (2000).
[12]R. R. Goruganthu and R. A. Bonham, Phys. Rev. A **26**, 1 (1982).
[13]R. Gaudoin and J. M. Pitarke, Phys. Rev. Lett. **99**, 126406 (2007).
[14]C. J. Umrigar, M. P. Nightingale, and K. J. Runge, J. Chem. Phys. **99**, 2865 (1993).