Bilinear diffusion quantum Monte Carlo methods

F. Arias de Saavedra*

Department of Chemistry, University of California, Berkeley, California 94720 and Lawrence Livermore National Laboratory, Livermore, California 94550

M. H. Kalos

Lawrence Livermore National Laboratory, Livermore, California 94550 (Received 19 April 2002; published 27 February 2003)

The standard method of quantum Monte Carlo for the solution of the Schrödinger equation in configuration space can be described quite generally as devising a random walk that generates—at least asymptotically populations of random walkers whose probability density is proportional to the wave function of the system being studied. While, in principle, the energy eigenvalue of the Hamiltonian can be calculated with high accuracy, estimators of operators that do not commute the Hamiltonian cannot. Bilinear quantum Monte Carlo (BQMC) is an alternative in which the square of the wave function is sampled in a somewhat indirect way. More specifically, one uses a pair of walkers at positions x and y and introduces stochastic dynamics to sample $\phi_i(x)t(x,y)\phi_j(y)$, where $\phi_i(x)$ and $\phi_j(y)$ are eigenfunctions of (possibly different) Hamiltonians, and t(x,y)is a kernel that correlates positions x and y. Using different Hamiltonians permits the accurate computation of small energy differences. We review the conceptual basis of BQMC, discuss qualitatively and analytically the problem of the fluctuations in the branching, and present partial solutions to that problem. Finally we exhibit numerical results for some model systems including harmonic oscillators and the hydrogen and helium atoms. Further research will be necessary to make this a practical and generally applicable scheme.

DOI: 10.1103/PhysRevE.67.026708

PACS number(s): 02.70.Ss, 02.70.Tt, 02.50.Ng

I. INTRODUCTION

The term quantum Monte Carlo (QMC) is used to encompass different approaches for solving the Schödinger equation of a system using stochastic techniques [1]. Most of these techniques represent the state in question (usually the ground state) as a set of positions in the space, which are distributed in proportion to the wave function for that state. This representation combined with a trial wave function can be used to calculate exactly for bosonic systems, both the energy of the state and expectations of operators that commute with the Hamiltonian. When we sum the values of the trial wave function over the positions of the random walk, we say that we are using a "mixed estimator."

QMC techniques have been applied with great success in atomic and molecular physics [1,2], nuclear physics [3], and condensed matter physics [4]. However, there are still open problems such as the sign problem for fermionic systems and the unbiased evaluation of the expectation value of operators that do not commute with the Hamiltonian. The sign problem is caused by the fact that the states of a fermionic system do not have a defined sign, so they cannot be used as the probability distribution function need by QMC techniques. We can eliminate this problem by multiplying by a trial function, but in this case we will not obtain the exact energy since we are forcing our representation to have the same zeros as the trial function. Different techniques have been applied to overcome this problem [5] and in the recent years the fermion Monte Carlo technique is providing very promising results [6].

In this paper, we present work in progress to find unbiased expectation values of operators that do not commute with the Hamiltonian. The fundamental difficulty is that most existing methods in QMC use a delta-function basis, and therefore the solution that arises from the random walk cannot be squared directly: it becomes necessary to "project" by way of a mixed estimator, which is simply the sum of a trial function over the positions generated in the random walk. BQMC avoids this dilemma by sampling the square of the wave function, albeit in an indirect way.

The mixed estimator is an approximation that can be improved by using the extrapolation method [7]. If ϕ_T is an approximation to the exact ground state ϕ of the system under study, one can approximate a correct estimator

$$A_{p} = \frac{\langle \phi | A | \phi \rangle}{\langle \phi | \phi \rangle} \tag{1}$$

as the linear combination of the mixed and variational estimators, respectively,

$$A_{m} = \frac{\langle \phi_{T} | A | \phi \rangle}{\langle \phi_{T} | \phi \rangle}, \quad A_{v} = \frac{\langle \phi_{T} | A | \phi_{T} \rangle}{\langle \phi_{T} | \phi_{T} \rangle}, \tag{2}$$

which has error quadratic in the difference between the two wave functions ($\Delta = \phi - \phi_T$). That is

$$A_p = 2A_m - A_v + O(\Delta^2). \tag{3}$$

This method has been widely used [8] because of its simplicity, but it provides an estimate of the expectation value whose bias is hard to assess. This problem of the bias has been eliminated by using two related methods: the future walking method and the time correlation method. The first one is based on the estimation of the ratio ϕ/ϕ_T using the

^{*}Permanent address: Departamento de Física Moderna, Facultad de Ciencias, Granada E-18071, Spain.

asymptotic offspring coming from the branching term [9]. In this line, different tagging algorithms have been constructed to properly account for the asymptotic number of descendants [10–12]. The tagging process can be eliminated by an evaluation of a weight proportional to the expected future progeny of every walker after each step [13]. This method is asymptotically exact and has been successfully applied to certain problems to obtain very accurate results [10–14]. However all of these methods, which rely on forward walking, are technically delicate and lead to a signal-to-noise ratio that decays to zero at large imaginary time, when the methods are least biased. The time correlation method also permits calculating the ratio ϕ/ϕ_T and has been recently applied in the context of path integral Monte Carlo [15].

Another challenging problem for QMC is the calculation of the energy difference between two similar systems. This can be expressed in terms of matrix elements of the wave functions. That is, if ϕ_i for i = 1,2 denote the ground states of the systems with Hamiltonians H_i , then

$$E_2 - E_1 = \frac{\langle \phi_1 | H_2 - H_1 | \phi_2 \rangle}{\langle \phi_1 | \phi_2 \rangle},\tag{4}$$

assuming that the number of degrees of freedom of both systems is the same. If an independent calculation of both energies is used for calculating the difference, the statistical variance of the mean can be comparable with the energy difference leading to no valuable information. For some systems, it has been possible to correlate the two random walks such that there is a important cancellation in the variance. However, this is possible only for few systems [16,17].

A solution to both problems of calculating energy differences and expectations of operators is provided by bilinear quantum Monte Carlo (BQMC). The basic idea is to correlate the configurations representing the two related wave functions. That is, BQMC aims to sample

$$\Psi(\vec{X}) = \phi_1(\vec{x})t(\vec{x}, \vec{y})\phi_2(\vec{y}).$$
 (5)

We shall hereafter use capital letters to denote pairs of vectors. The function $t(\vec{x}, \vec{y})$ is chosen to correlate the configurations \vec{x} and \vec{y} . Thus the elements in the random walk are pairs of configurations $\vec{X} = (\vec{x}, \vec{y})$ rather than individual configurations. The basic idea is that if we choose $t(\vec{x}, \vec{y})$ equal to Green's function for the second system, then, since

$$\int d\vec{y} \phi_1(\vec{x}) G_2(\vec{x}, \vec{y}) \phi_2(\vec{y}) = \phi_1(\vec{x}) \phi_2(\vec{x}), \quad (6)$$

the marginal distribution samples the product of the wave functions. This will allow us to compute energy differences and, in the case when the two function are the same, to calculate correct expectations of general operators.

This method was first proposed by Zhang and Kalos [18] and developed in the framework of Green's function Monte Carlo. They applied it to several test problems including a model integral equation and the hydrogen atom and obtained good results. However, the technical problem of large fluctuations in the number of pairs of walkers in a MC step was not completely solved by the authors. We have analyzed and eliminated this problem and have extended BQMC by using the imaginary-time-dependent Green's function formalism, that is, within the framework of diffusion Monte Carlo (DMC). We believe that this will be more accessible, and will lead more easily to practical realizations of BQMC.

The structure of this paper is the following. In Sec. II, we present BQMC in more detail and construct the different estimators used in this work. After that, we study as a first application the case of the harmonic oscillator in any number of dimensions. This is a first contact with the characteristics of the method when no short-time approximation is needed for Green's function of the system, since it is known exactly [19]. In Sec. IV, we present the expression for BQMC in the short-time limit; this generates a set of equations with a similar structure to that of DMC when importance sampling is used. The application of these equations for the harmonic oscillator and the hydrogen and helium atoms will be covered in Sec. V. Finally we offer our conclusions in Sec. VI.

II. DESCRIPTION OF THE METHOD

We have remarked above that the bilinear Quantum Monte Carlo method can be applied in two different situations. The first one is to calculate the expectation value of any operator, i.e., to sample the square of the wave function. The second one is to calculate the difference in energies between two very similar systems. In the following discussion we will adopt this last case since the first one can be recovered considering that the two wave functions are the same and choosing $t(\vec{x}, \vec{y})$ appropriately. We introduce integral equations satisfied by the functions

$$\Psi(\vec{X},\tau_1) = \phi_1(\vec{x},\tau_1)t(\vec{x},\vec{y})\phi_2(\vec{y},\tau_1)$$
(7)

that have the form

$$\Psi(\vec{X},\tau_1) = \lambda \int d\vec{U} G(\vec{X},\vec{U};\tau_1) \Psi(\vec{U},0).$$
(8)

The elements involved in this random walk are pairs of walkers. If we take into account that

$$\phi_{i}(\vec{x},\tau_{1}) = \lambda_{i} \int d\vec{u} G_{i}(\vec{x},\vec{u};\tau_{1}) \phi_{i}(\vec{u},0), \qquad (9)$$

we easily obtain

$$G(\vec{X}, \vec{U}; \tau_1) = \frac{G_1(\vec{x}, \vec{u}; \tau_1) t(\vec{x}, \vec{y}) G_2(\vec{y}, \vec{v}; \tau_1)}{t(\vec{u}, \vec{v})}$$
(10)

and $\lambda = \lambda_1 \lambda_2$. We can rewrite Eq. (8) to make it more transparent for a random walk interpretation:

$$\Psi(\vec{X},\tau_1) = \lambda \int d\vec{U} \ \Gamma(\vec{X}|\vec{U};\tau_1) N(\vec{U},\tau_1) \Psi(\vec{U},0), \ (11)$$

 $N(\tilde{U}, \tau_1)$ gives the multiplicity of walkers that come from the initial pair \vec{U} ,

$$N(\vec{U},\tau_1) = \int d\vec{X} G(\vec{X},\vec{U};\tau_1) \tag{12}$$

is the branching factor. Then Γ is a normalized probability density function of the pair \vec{X} conditional on the pair \vec{U} ,

$$\Gamma(\vec{X}|\vec{U};\tau_1) = \frac{G(\vec{X},\vec{U};\tau_1)}{N(\vec{U},\tau_1)}.$$
(13)

The random walk involves branching of the pairs after which they are sampled according to Γ . Of course, we must be able to carry out in a practical way the integration in the definition of *N* and to sample Γ .

We now discuss the form of the estimators, beginning with the energy difference. A correct estimator for this using any choice of $t(\vec{x}, \vec{y})$ is

$$E_{2}-E_{1} = \frac{\int d\vec{x}d\vec{y}\phi_{1}(\vec{x})t(\vec{x},\vec{y})H_{2}(\vec{y})\phi_{2}(\vec{y})}{\int d\vec{x}d\vec{y}\phi_{1}(\vec{x})t(\vec{x},\vec{y})\phi_{2}(\vec{y})} - \frac{\int d\vec{x}d\vec{y}\phi_{1}(\vec{x})H_{1}(\vec{x})t(\vec{x},\vec{y})\phi_{2}(\vec{y})}{\int d\vec{x}d\vec{y}\phi_{1}(\vec{x})t(\vec{x},\vec{y})\phi_{2}(\vec{y})} = \frac{\int d\vec{x}\Psi(\vec{X})[E_{2,L}(\vec{y}) - E_{1,L}(\vec{x})]}{\int d\vec{x}\Psi(\vec{X})}, \quad (14)$$

where for Hamiltonians of the form

$$H_i(\vec{x}) = -\frac{1}{2m_i} \nabla_{\vec{x}}^2 + V_i(\vec{x}), \qquad (15)$$

we define two different local energies as

$$E_{i,L}(\vec{a}) = -\frac{\nabla_{\vec{a}}^2 t(\vec{x}, \vec{y})}{2m_i t(\vec{x}, \vec{y})} + V_i(\vec{a}).$$
(16)

In the case when $m_1 = m_2$, we can build a better estimator of the difference of energies by choosing $t(\vec{x}, \vec{y})$ appropriately. That is, if

$$t(\vec{x}, \vec{y}) = G_2(\vec{x}, \vec{y}; \tau_2), \tag{17}$$

then

$$d\vec{y}\Psi(\vec{X}) = \exp(\tau_2 E_2)\phi_1(\vec{x}) \cdot \phi_2(\vec{x}).$$
 (18)

In this case

$$E_{2}-E_{1} = \frac{\int d\vec{X}\Psi(\vec{X})(E_{2,L}(\vec{x}) - E_{1,L}(\vec{x}))}{\int d\vec{X}\Psi(\vec{X})}$$
$$= \frac{\int d\vec{X}\Psi(\vec{X})(V_{2}(\vec{x}) - V_{1}(\vec{x}))}{\int d\vec{X}\Psi(\vec{X})}.$$
(19)

This choice of $t(\vec{x}, \vec{y})$ is also the one that must be used when both systems are the same and we are calculating expectation values of general operators. If $\phi_2 = \phi_1$ we can compute the expectation value of any function of the spatial coordinates, $f(\vec{y})$, by using

$$\langle f \rangle = \frac{\int d\vec{x} d\vec{y} \phi_1(\vec{x}) f(\vec{x}) G_1(\vec{x}, \vec{y}; \tau_2) \phi_1(\vec{y})}{\int d\vec{x} d\vec{y} \phi_1(\vec{x}) G_1(\vec{x}, \vec{y}; \tau_2) \phi_1(\vec{y})} = \frac{\int d\vec{x} \phi_1(\vec{x}) f(\vec{x}) \phi_1(\vec{x})}{\int d\vec{x} \phi_1(\vec{x}) \phi_1(\vec{x})}.$$
 (20)

This can be easily generalized to any operator. In this case

$$\langle O \rangle = \frac{\int d\vec{x} d\vec{y} \phi_1(\vec{x}) O(\vec{x}) G_1(\vec{x}, \vec{y}; \tau_2) \phi_1(\vec{y})}{\int d\vec{x} d\vec{y} \phi_1(\vec{x}) G_1(\vec{x}, \vec{y}; \tau_2) \phi_1(\vec{y})} = \frac{\int d\vec{X} \Psi(\vec{X}) O_L(\vec{x})}{\int d\vec{X} \Psi(\vec{X})},$$
(21)

with the usual definition

$$O_L(\vec{x}) = \frac{O(\vec{x})G_1(\vec{x}, \vec{y}; \tau_2)}{G_1(\vec{x}, \vec{y}; \tau_2)}.$$
 (22)

After some relaxation time, our representation of the density of walkers bilinear in the wave functions becomes

$$\Psi(\vec{X}) = \phi_1(\vec{x})t(\vec{x},\vec{y})\phi_2(\vec{y}) \rightarrow \{\vec{X}_k\}_{k=1}^M = \{(\vec{x}_k,\vec{y}_k)\}_{k=1}^M,$$
(23)

with M being the total number of pairs of walkers.

In order to estimate the difference of energies between the systems, we use

$$E_2 - E_1 = \frac{1}{M} \sum_{k=1}^{M} [E_{2,L}(\vec{y}_k) - E_{1,L}(\vec{x}_k)].$$
(24)

Setting $m_1 = m_2$ and choosing $t(\vec{x}, \vec{y}) = G_2(\vec{x}, \vec{y}; \tau_2)$, we obtain

$$E_2 - E_1 = \frac{1}{M} \sum_{k=1}^{M} \left[V_2(\vec{x}_k) - V_1(\vec{x}_k) \right].$$
(25)

This equation is formally identical to the integral Hellmann-Feynman theorem [20] and this choice of $t(\vec{x}, \vec{y})$, when both systems are the same, allows us to write for any operator

$$\langle O \rangle = \frac{1}{M} \sum_{k=1}^{M} O_L(\vec{y}_k) = \frac{1}{M} \sum_{k=1}^{M} O_L(\vec{x}_k),$$
 (26)

since both marginal distributions are identical and are distributed as the square of the ground state wave function of the system.

Bilinear diffusion Monte Carlo (BDMC) method may suffer from its own forms of bias in addition to the usual timestep, relaxation, and population-control errors of ordinary (linear) diffusion Monte Carlo. A good opportunity for exploring and eliminating these problems would be a system whose exact time-dependent Green's function is known. Then we would not have to concern ourselves about the magnitude of the imaginary time since there is no finite timestep error. Green's function for the harmonic oscillator is known analytically and easily sampled for any number of degrees of freedom: this will be our first application.

III. BILINEAR MONTE CARLO FOR THE HARMONIC OSCILLATOR

The imaginary-time Schrödinger equation for an *N*-dimensional harmonic oscillator is

$$-\hbar \frac{\partial \phi(\vec{x},\tau)}{\partial \tau} = \left(E + \frac{\hbar^2}{2m} \nabla_{\vec{x}}^2 - \frac{1}{2} m \omega \vec{x}^2 \right) \phi(\vec{x},\tau). \quad (27)$$

This Hamiltonian depends on the mass m and the oscillator constant ω . Introducing reduced coordinates for the energy, time, and position we obtain

$$\vec{x}' = \sqrt{\frac{m\omega}{\hbar}}\vec{x}, \quad E' = \frac{E}{\hbar\omega}, \quad \tau' = \omega\tau,$$
 (28)

the result is

$$-\frac{\partial \phi(\vec{x}', \tau')}{\partial \tau'} = \left(E' + \frac{1}{2}\nabla_{\vec{x}'}^2 - \frac{1}{2}\vec{x}'^2\right)\phi(\vec{x}', \tau').$$
 (29)

This is equivalent to making $\hbar = 1$ and $m = \omega = 1$ in Eq. (27). Thus when we work with only one harmonic oscillator we can choose $m = \omega = 1$. With different harmonic oscillators one of them can have $m_1 = \omega_1 = 1$, and the second

$$H_{m,w}(\vec{x}) = -\frac{1}{2m}\nabla_{\vec{x}}^2 + \frac{1}{2}m\omega\vec{x}^2$$
(30)

and we can set $\omega \ge 1$. As noted above, the time-dependent Green's function is known for the harmonic oscillator [19]:

$$G_{m,\omega}(\vec{x}, \vec{y}; \tau) = \left(\frac{m\omega}{2\pi\sinh\omega\tau}\right)^{N/2} \\ \times \exp\left(-\frac{m\omega}{2\tanh\omega\tau}(\vec{x}^2 + \vec{y}^2) + \frac{m\omega}{\sinh\omega\tau}\vec{x}\cdot\vec{y}\right).$$
(31)

We now analyze the case with one Hamiltonian. Here, the branching and sampling parts of Green's equation for the bilinear wave function are, respectively,

$$N(\vec{u}, \vec{v}) = \exp(2E\tau_1) \frac{G_{1,1}(\vec{u}, \vec{v}, 2\tau_1 + \tau_2)}{G_{1,1}(\vec{u}, \vec{v}, \tau_2)}, \qquad (32)$$

$$\Gamma(\vec{x}, \vec{y} | \vec{u}, \vec{v}) = \left(\frac{\sqrt{\alpha\beta}}{2\pi}\right)^N \exp\left(-\frac{\alpha}{2}(\vec{x} - \vec{x}_0)^2 - \frac{\beta}{2}(\vec{y} - \vec{y}_0)^2\right),\tag{33}$$

with

$$\alpha = \frac{\sinh(\tau_1 + \tau_2)}{\sinh\tau_1 \sinh\tau_2},\tag{34}$$

$$\vec{x}_0 = \frac{\vec{u}}{\alpha \sinh \tau_1} + \frac{\vec{y}}{\alpha \sinh \tau_2},$$
(35)

$$\beta = \frac{\sinh(2\tau_1 + \tau_2)}{\sinh\tau_1\sinh(\tau_1 + \tau_2)},\tag{36}$$

$$\vec{y}_0 = \frac{\vec{v}}{\beta \sinh \tau_1} + \frac{\vec{u}}{\beta \sinh(\tau_1 + \tau_2)}.$$
(37)

Now we study the behavior of our systems in terms of the two imaginary times, τ_1 and τ_2 from Eqs. (7) and (17). However, we focus on small values of τ_1 since this time controls the evolution of the pairs of walkers and we are interested in applying this method to more general systems, when we will use the short-time limit in order to get an easy approximate time-dependent Green's function. In the present case there is no limitation in the values of the times that can be used and we will explore this possibility for τ_2 .

We have not considered values of τ_2 smaller than τ_1 , since in this case the fluctuations of the branching factor can be very large; its moments may not be finite. As a matter of fact, the condition

$$\tanh(2\tau_1 + \tau_2) < \frac{n \tanh \tau_2}{n - 1 - \tanh \tau_2} \tag{38}$$

must be satisfied in order that the *n*th moment of the branching factor be finite. The definitions of these moments and further details are given in the Appendix. There we also discuss how the following short-time limit condition for τ_1 and τ_2 is obtained:

$$\tau_2 \ge 2(n-1)\tau_1,\tag{39}$$

that is, if τ_2 were smaller than τ_1 , even the mean square of N would diverge.



FIG. 1. Comparison of the potential (top) and total (bottom) energies with the exact values as a function of τ_2 for two values of τ_1 . The diamonds correspond to $\tau_1 = 0.001$ and the crosses to $\tau_1 = 0.01$.

For simplicity, we begin considering the one-dimensional case (N=1) with 4000 walkers and we study different values of τ_2 for $\tau_1 = 0.001$ and 0.01. The results obtained for the potential and total energy are shown in the Fig. 1, as a function of τ_2 . We can see that there are significant deviations from the exact values for values of τ_2 smaller than 0.5 and that the deviations are larger for the total energy than for the potential energy. This behavior is caused by a correlation between the members of the pair that causes substantial fluctuations in the population. The similar behavior has been found in the three-dimensional system.

The applicability of the method more generally requires the relaxation of this condition to small τ_2 . A possible solution to this problem is to use a portion of Green's function that correlates the pair as a weight and sample the rest. That is, set

$$t(\vec{x}, \vec{y}) = [G_{1,1}(\vec{x}, \vec{y}; \tau_2)]^{\mu}$$
(40)

with μ an exponent between 0 and 1. In this case and in the short-time limit the condition (39) becomes

$$\frac{\tau_2}{\mu} \! \ge \! 2(n\!-\!1)\,\tau_1, \tag{41}$$

so this choice of $t(\vec{x}, \vec{y})$ is equivalent to using a time τ_2/μ instead of τ_2 . This makes the fluctuations of the branching





FIG. 2. Comparison of the potential and total energies for two choices of τ_1 ($\tau_2 = 5 \tau_1$) as a function of μ . The upper panel corresponds to $\tau_1 = 0.001$ and the lower panel corresponds to $\tau_1 = 0.01$. The diamonds represent the total energy and the crosses represent the potential energy.

factor disappear for a small enough value of μ . Broadening the Green's function can be interpreted as a particular case of importance sampling when we use $G^{\mu-1}$ as importance function. Since we can use any function between the pair and get the exact value of the energy, we shall use as energy estimator

$$E = \frac{1}{M} \sum_{k=1}^{M} \left(-\frac{\nabla_{\vec{x}}^2 [G_{1,1}(\vec{x}_k, \vec{y}_k; \tau_2)]^{\mu}}{2 [G_{1,1}(\vec{x}_k, \vec{y}_k; \tau_2)]^{\mu}} + \frac{1}{2} \vec{x}_k^2 \right).$$
(42)

In order to get an unbiased answer, we must divide by the importance function. Hence the expectation value of the potential is

$$\langle V \rangle = \frac{\sum_{k=1}^{M} \frac{\vec{x}_{k}^{2}}{2} [G_{1,1}(\vec{x}_{k}, \vec{y}_{k}; \tau_{2})]^{1-\mu}}{\sum_{k=1}^{M} [G_{1,1}(\vec{x}_{k}, \vec{y}_{k}; \tau_{2})]^{1-\mu}}.$$
(43)

Computations of the total and the potential energy were carried out varying the parameter μ in two different cases for $\tau_1 = 0.001, 0.01$ and $\tau_2 = 5\tau_1$. We chose these values since there are important biases in both quantities for $\mu = 1$ and we can consider that they both pertain to the short-time limit. The results are shown in Fig. 2. We can see that the bias

TABLE I. Potential and total energies for different choices of τ_2 of the one-dimensional harmonic oscillator where $\tau_1 = 0.001$ and $\mu = \tau_2/2$.

$ au_2$	$\langle V \rangle$	Е
0.001	0.25002(39)	0.50011(37)
0.002	0.24979(31)	0.49976(32)
0.005	0.25010(42)	0.50016(44)
0.010	0.24998(33)	0.50005(35)
0.100	0.24982(32)	0.50009(34)

disappears markedly for $\mu \leq \tau_2$, indicating that the elimination of the bias in the results is related to τ_2/μ rather than to τ_2 . The same behavior is present in a three-dimensional calculation.

Results for $\tau_1 = 0.001$, different choices of τ_2 and μ $= \tau_2/2$ are given in Table I. The number of walkers is 40 000, and there are 20 blocks with 10 000 steps per block. All the choices of τ_2 (even $\tau_2 = \tau_1$) provide unbiased values and the values of the variance do not change significantly with the imaginary time. Therefore the introduction of a suitable choice of $t(\vec{x}, \vec{y})$ in the density of walkers bilinear in the wave function can give unbiased values for the estimators for any choice of τ_1 and τ_2 with $\tau_2 \ge \tau_1$. This result is very important since it allows us to work in the short-time limit for both imaginary times and indicates that we can use standard approximations for time-dependent Green's functions, making the method applicable to more general systems. Now we consider the behavior of BDMC for two different harmonic oscillators. As already discussed, we choose the first with $m = \omega = 1$ and for the second one we vary m and ω . We first used $t(\vec{x}, \vec{y}) = G_{1,1}(\vec{x}, \vec{y}; \tau_2)$ and found the same behavior as before. That is, in order to obtain an unbiased result for the energy difference τ_2 must be large compared to τ_1 . A good choice for eliminating the bias is again t(x,y) $= [G_{1,1}(\vec{x}, \vec{y}; \tau_2)]^{\mu}$ and $\mu \ge \tau_2$. Again the estimator used in the case m=1 must be weighted since the whole Green's function must be accounted for; that is,

$$E_{2}-E_{1} = \frac{\frac{\omega-1}{2} \sum_{k=1}^{M} \vec{y}_{k}^{2} [G_{1,1}(\vec{x}_{k}, \vec{y}_{k}; \tau_{2})]^{1-\mu}}{\sum_{k=1}^{M} [G_{1,1}(\vec{x}_{k}, \vec{y}_{k}; \tau_{2})]^{1-\mu}}.$$
 (44)

Table II exhibits the energy difference between two onedimensional harmonic oscillators results for different values of ω and m, using $\tau_1 = 0.001$, $\tau_2 = 0.01$, and $\mu = \tau_2/2$. The number of walkers, the size of the block and the number of the blocks are the same as that for the calculations presented in the Table I. The energy difference estimator depends on whether m=1 or $m \neq 1$. We can see that there are good results with variances even smaller than those for the energy. This reduction of the variance is especially important in the case where the two systems have equal masses as a consequence of our choice of the estimator, as may be seen by

TABLE II. Energy difference between two one-dimensional harmonic oscillators with masses m and 1 and oscillator constants ω and 1.

т	ω	$E_2 - E_1$	Exact
1.0	1.0100	0.0050045(76)	0.005
1.0	1.0001	0.000050025(72)	0.00005
1.1	1.0000	-0.00003(18)	0
2.0	1.0000	-0.00002(26)	0

comparing the first row of the table with $0.005\,12(27)$, the result obtained from the other estimator.

We are now ready to extend the bilinear diffusion method to systems with unknown time-dependent Green's function, using two small time steps so that easy approximations can be obtained for them.

IV. BILINEAR DIFFUSION MONTE CARLO IN THE SHORT-TIME LIMIT

It is known [21,22] that in the short-time limit an approximation to the time-dependent Green's functions of the systems with Hamiltonians H_i has the form

$$G_{i}(\vec{x}, \vec{y}; \tau) = \left(\frac{m_{i}}{2 \pi \tau}\right)^{N/2} \exp\left(-\frac{m_{i}}{2 \tau}(\vec{x} - \vec{y})^{2}\right)$$
$$\times \exp\left(-\frac{\tau}{2}\left[V_{i}(\vec{x}) + V_{i}(\vec{y})\right]\right).$$
(45)

If we choose a form for $t(\vec{x}, \vec{y})$, we can perform the same analysis as in the case where the exact time-dependent Green's function is known. Instead of that, we analyze the time evolution of our density of walkers, $\Psi(\vec{X}, \tau_1)$, defined in Eq. (5). The time-dependent Schrödinger equations for the two systems are

$$\frac{\partial \phi_i(\vec{x},\tau)}{\partial \tau} = [E_i - H_i(\vec{x})] \phi_i(\vec{x},\tau), \quad i = 1,2.$$
(46)

The time-dependent equation satisfied by Ψ is then

$$\frac{\partial \Psi(\vec{X},\tau_{1})}{\partial \tau_{1}} = t(\vec{x},\vec{y}) \frac{\partial \phi_{1}(\vec{x},\tau_{1})}{\partial \tau_{1}} \phi_{2}(\vec{y},\tau_{1}) \\
+ t(\vec{x},\vec{y}) \phi_{1}(\vec{x},\tau_{1}) \frac{\partial \phi_{2}(\vec{y},\tau_{1})}{\partial \tau_{1}} \\
= t(\vec{x},\vec{y}) [E_{1} + E_{2} - H_{1}(\vec{x}) - H_{2}(\vec{y})] \\
\times \phi_{1}(\vec{x},\tau_{1}) \phi_{2}(\vec{y},\tau_{1}) \\
= t(\vec{x},\vec{y}) [E_{1} + E_{2} - H_{1}(\vec{x}) - H_{2}(\vec{y})] \\
\times \left(\frac{\Psi(\vec{X},\tau_{1})}{t(\vec{x},\vec{y})}\right), \quad (47)$$

which has the same structure as the importance-sampled Schrödinger equation in imaginary time. We define the quantum forces in the usual way

$$\vec{F}_{\vec{a}} = \frac{2\vec{\nabla}_{\vec{a}}t(\vec{x},\vec{y})}{t(\vec{x},\vec{y})},$$
(48)

where \vec{a} is either \vec{x} or \vec{y} ; we then have

$$\frac{\partial \Psi(X,\tau_{1})}{\partial \tau_{1}} = (E_{1} + E_{2} - H_{1}(\vec{x}) - H_{2}(\vec{y}))\Psi(\vec{X},\tau_{1}) \\ + \left(\frac{\nabla_{\vec{x}}^{2}t(\vec{x},\vec{y})}{2m_{1}t(\vec{x},\vec{y})} + \frac{\nabla_{\vec{y}}^{2}t(\vec{x},\vec{y})}{2m_{2}t(\vec{x},\vec{y})}\right)\Psi(\vec{X},\tau_{1}) \\ - \frac{1}{2m_{1}}\vec{\nabla_{\vec{x}}} \cdot [\Psi(\vec{X},\tau_{1})\vec{F}_{\vec{x}}] \\ - \frac{1}{2m_{2}}\vec{\nabla_{\vec{y}}} \cdot [\Psi(\vec{X},\tau_{1})\vec{F}_{\vec{y}}].$$
(49)

Making the usual analysis [21], we can write in the short-time limit

$$\Gamma(\vec{X}|\vec{U};\tau_{1}) = \left(\frac{m_{1}}{2\pi\tau_{1}}\right)^{N/2} \exp\left[-\frac{m_{1}}{2\tau_{1}}\left(\vec{x}-\vec{u}-\frac{\tau_{1}}{2m_{1}}\vec{F}_{\vec{u}}\right)^{2}\right] \\ \times \left(\frac{m_{2}}{2\pi\tau_{1}}\right)^{N/2} \exp\left[-\frac{m_{2}}{2\tau_{1}}\left(\vec{y}-\vec{v}-\frac{\tau_{1}}{2m_{2}}\vec{F}_{\vec{v}}\right)^{2}\right],$$
(50)

$$N(\vec{U},\tau_1) = \exp[\{E_1 + E_2 - E_{1,L}(\vec{u}) - E_{2,L}(\vec{v})\}\tau_1], \quad (51)$$

with $\vec{X} = (\vec{x}, \vec{y})$ and $\vec{U} = (\vec{u}, \vec{v})$. The time-dependent Green's functions is separated into its sampling and branching parts. The sampling part has the drift terms that depend on both \vec{u} and \vec{v} , reflecting the correlation of the walkers.

In general, we have used

$$t(\vec{x}, \vec{y}) = [G_1(\vec{x}, \vec{y}; \tau_2)]^{\mu}$$

= $\left(\frac{m_1}{2\pi\tau_2}\right)^{\mu N/2} \exp\left(-\frac{\mu m_1}{2\tau_2}(\vec{x} - \vec{y})^2\right)$
 $\times \exp\left(-\frac{\mu \tau_2}{2}[V_1(\vec{x}) + V_1(\vec{y})]\right),$ (52)

TABLE III. Comparison between the potential and total energies for different choices of μ of the one-dimensional harmonic oscillator when the exact propagator results are compared with the diffusion results.

$\langle V \rangle$			E	
μ	Exact	Diff.	Exact	Diff.
0.010	0.24965(22)	0.25009(30)	0.49975(44)	0.50014(55)
0.005	0.24998(33)	0.24999(37)	0.50005(35)	0.49987(33)

TABLE IV. Comparison of the energy differences between two one-dimensional harmonic oscillators with masses m and 1 and oscillator constants ω and 1 for the cases when the exact time-dependent Green's function is used and when the diffusion equation is used.

		$E_2 - E_1$		
т	ω	Exact	Diff.	
1.0	1.0100	0.0050045(76)	0.0050083(84)	
1.0	1.0001	0.000050025(72)	0.000050031(78)	
1.1	1.0000	-0.00003(18)	0.00021(18)	
2.0	1.0000	-0.00002(26)	-0.00003(32)	

where μ is the chosen exponent for the Green's function, between 0 and 1, that allows the elements of a pair to be less tightly correlated.

V. APPLICATIONS IN THE SHORT-TIME LIMIT

We have applied our method to three different systems: harmonic oscillators and the hydrogen and helium atoms, repeating the first to check whether additional problems arise when the exact propagator is replaced by its short-time approximation.

First we compare the results obtained for the potential and total energies of the one-dimensional harmonic oscillator when approximate propagators are used. Table III shows results for $\tau_1 = 0.001$ and $\tau_2 = 0.01$. We can see that there is no bias in the diffusion results and that the variance is almost the same as that obtained using the exact propagator.

Next we made BDMC calculations for all the cases of the energy differences presented in Table II. The comparison with these data is shown in Table IV. Again we observe that there are no significant differences between using the exact propagator and its short-time approximations.

We are now in a position to study a more realistic system: the hydrogen atom with the infinite nuclear mass approximation. This is a tridimensional system whose main difference from the harmonic oscillator is the singularity in the potential at the nucleus. When this system is studied with DMC, an importance sampling function that satisfies Kato's cusp condition [24] must be used to eliminate the singularity. However, we first present in Table V results computed without an importance function, using $\tau_1 = 0.001$, $\tau_2 = 0.01$, and μ

TABLE V. Comparison for different expectation values in terms of the number of pairs of walkers N_w with the exact result for the hydrogen atom.

N _w	Ε	$\langle -1/r \rangle$	$\langle r \rangle$	$\langle r^2 \rangle$
4000	-0.4935(43)	-0.9929(69)	1.4986(81)	2.972(32)
8000	-0.4988(25)	-0.9992(50)	1.4970(55)	2.978(23)
16000	-0.4960(20)	-1.0021(38)	1.4940(48)	2.969(22)
32000	-0.5015(14)	-1.0017(22)	1.4937(28)	2.966(12)
64000	-0.49893(79)	-1.0005(20)	1.4995(26)	2.997(13)
Exact	-0.5	-1	1.5	3



FIG. 3. Density of the probability of the hydrogen atom for the calculation of 64 000 walkers shown in Table V.

 $=\tau_2/2$. They comprise 20 blocks of 10 000 steps each. We can see that the results show no bias. In the case of 64 000 walkers, we show in Fig. 3 a histogram of the density of walkers; there is no appreciable difference from the exact result:

$$D(r) = 4r^2 \exp(-2r).$$
 (53)

Our next refinement was to use an importance function:

$$\phi_T(\vec{r}) = \exp\left(-\frac{\alpha r}{\alpha + r}\right). \tag{54}$$

For any choice of the parameter, α , the cusp condition is satisfied and moreover

$$\lim_{\alpha \to \infty} \phi_T(\vec{r}) = \exp(-r) = \phi(\vec{r}).$$
(55)

This function is included by replacing Eq. (52) by

$$t(\vec{x}, \vec{y}) = \phi_T(\vec{x}) G_1^{\mu}(\vec{x}, \vec{y}; \tau_2) \phi_T(\vec{y})$$
(56)

in the definitions of the quantum forces and local energies. This has the correct effect of biasing the random walk to reflect the effect of $\phi_T(\vec{r})$ and removing the bias in calculating average energies.

TABLE VI. Comparison for different expectation values in terms of the number of pairs of walkers N_w with the exact result for the hydrogen atom using the importance sampling discussed in the text.

N_w	Ε	$\langle -1/r \rangle$	$\langle r \rangle$	$\langle r^2 \rangle$
4000	-0.4952(32)	-0.9967(70)	1.4968(65)	2.975(25)
8000	-0.4968(18)	-0.9998(45)	1.4865(49)	2.933(19)
16000	-0.4991(14)	-0.9970(50)	1.5005(57)	2.993(18)
32000	-0.5028(27)	-0.9930(62)	1.516(16)	3.083(83)
64000	-0.49893(94)	-0.9990(24)	1.5001(36)	2.998(16)
Exact	-0.5	-1	1.5	3

We have done calculations for the same conditions as in Table V for $\alpha = 0.1$; the results are shown in Table VI. We can see that the results are very similar to those obtained without importance sampling.

It is important to note that the role played by the importance sampling functions in BQMC is different from the familiar DMC case. For the latter, an importance function close to the exact wave function makes the local energy nearly constant, so that the fluctuations in the population are small. One samples a density close to the square of the wave function. On the other hand, BQMC already samples the square of the wave function without any importance function. If we use an importance function close to the exact one, we would sample the wave function to the fourth power. This would create large fluctuations when the wave function is small, say at large distances from the nucleus. Such inefficiency does not arise for our choice of importance sampling function since

$$\lim_{r \to \infty} \phi_T(r) = \exp(-\alpha), \tag{57}$$

so that far from the nucleus, the importance function does not vanish and does not generate any spurious fluctuations. The importance function serves only to remove any singularities in the local energy, and must therefore satisfy any cusp conditions.

Finally we study the ground state of the helium atom supposing again that the nucleus has an infinite mass. This is a six-dimensional system. Preliminary calculations using the usual choices of τ_1 , τ_2 , and μ , but without any importance function, yielded biased results as well as significant fluctuations in the population. The latter effect derives from the singularities in the local energy at the nucleus and disappeared when we introduced an importance function that satisfies the cusp conditions:

$$\phi_T(\vec{r}_1, \vec{r}_2) = \exp\left(-\frac{Z\alpha r_1}{\alpha + r_1} - \frac{Z\alpha r_2}{\alpha + r_2} + \frac{\beta r_{12}}{2(\beta + r_{12})}\right),$$
(58)

with Z being the charge on the nucleus. This is the obvious generalization of the form used for the hydrogen atom. After some experimentation we observed that the bilinear quantities were still biased even though there were no appreciable fluctuations of the population. It turned out that only a small fraction of the population carried significant weights in the evaluation of bilinear quantities. These weights are a rapidly varying function of the separation of the walkers in a pair, so that the fluctuations are sensitive to the distribution of separations, and therefore to the dimensionality. In other words, the bias is a consequence of inadequate importance sampling of the pair separations.

An obvious solution to this problem is to smooth the function that determines the pair separation. Since this function is related to the time-dependent Green's function with imaginary time τ_2 , the function is less peaked when a higher value of τ_2 is used. Within the short-time approximation, we must be careful with its magnitude. However, the role played here

TABLE VII. Comparison for different expectation values in terms of the number of pairs of walkers N_w with the exact result for the helium atom using the importance sampling discussed in the text.

N_w	Ε	$\langle V \rangle$	$\langle r_1 \rangle$	$\langle r_1^2 \rangle$
4000	-2.890(8)	-5.79(2)	0.929(5)	1.19(2)
8000	-2.896(3)	-5.78(1)	0.933(3)	1.202(9)
16000	-2.897(3)	-5.78(1)	0.932(2)	1.193(6)
32000	-2.900(4)	-5.804(7)	0.932(2)	1.195(5)
64000	-2.899(2)	-5.796(6)	0.933(4)	1.21(2)
128000	-2.900(2)	-5.807(2)	0.930(1)	1.191(3)
Exact	-2.9037	-5.8074	0.9294	1.1935

by τ_2 is not the same as the usual diffusion time step, which governs the time evolution of the walkers, for us, τ_1 .

Results obtained for $\tau_1 = 0.001$, $\tau_2 = 0.1$, and $\mu = 0.1$ and with $\alpha = \beta = 0.2$ for the importance sampling parameters are shown in Table VII. The exact results are taken from a precise variational calculation [23]. We can see that the BQMC results are quite good especially for the largest population. In this case we also calculated other expectation values, $\langle 1/r_{12} \rangle$, $\langle r_{12} \rangle$, and $\langle r_{12}^2 \rangle$. We have obtained, respectively, 0.9413(7), 1.424(1), and 2.521(5) to be compared to the exact ones 0.9458, 1.422, and 2.516. All these results, except the biased result for $\langle 1/r_{12} \rangle$, seem to indicate that a higher value for τ_2 does not generate any additional bias. In addition, the ratio of standard deviation to the average value of the weights is drastically improved with the larger value.

We also studied the behavior of different electron densities for the helium atom. The results obtained for the 128 000-walker calculation of the charge density, D(r), are compared to precise values from [23] in the top panel of Fig. 4. We note differences between both densities around the maximum that may indicate that this quantity is sensitive to the deviations of the time-dependent Green's function of the pair from its short-time limit. The sensitivity disappeared when we calculated the electron density as a function of the separation between the two electrons, the so-called intracule density, H(r). This is shown in the lower plot of Fig. 4.

In summary, the accurate results validate the BQMC approach. However, the technique is not as efficient as we would wish, due to the increased variance of the weights in the calculations of the helium atom. A more complete understanding of the source of inefficiency and its cure will be a subject of further research.

VI. CONCLUSIONS

BQMC (Bilinear quantum Monte Carlo) has allowed us to sample the square of the wave function and to calculate small energies differences efficiently for some simple systems. We have been able to control the fluctuations in the normalization or branching by an appropriate choice of the function that correlates the members of the pairs. We have adapted the use of short-time Green's functions to the bilinear case with a simple and transparent implementation and



FIG. 4. Comparison of the charge (top) and intracule (bottom) densities obtained with the BQMC (dashes lines) to the exact ones (solid lines).

applied it successfully to the harmonic oscillator and the hydrogen and helium atoms.

ACKNOWLEDGMENTS

This work was performed under the auspices of the U.S. Department of Energy by the University of California, Lawrence Livermore National Laboratory under Contract No. W-7405-Eng-48. F.A. thanks Professor W.A. Lester, Jr. of the Chemistry Department in the University of California at Berkeley, and Dr. Steven F. Ashby and Dr. Peter G. Elt-groth of CASC at Lawrence Livermore National Laboratory for their warm hospitality, and the Spanish Secretaría de Estado de Educación y Universidades for financial support. We are grateful for useful conversations with Professor Lester and Alán Aspuru-Guzik.

APPENDIX

The *n*th moment of the branching factor is defined as

$$M_n = \int d\vec{U} N^n(\vec{U}) \Psi(\vec{U}). \tag{A1}$$

If we use the expression for N(32) and take into account that

$$\Psi(\vec{U}) = \frac{1}{\sqrt{\pi}} \exp\left(-\frac{1}{2}(\vec{u}^2 + \vec{v}^2)\right) G_{1,1}(\vec{u}, \vec{v}, \tau_2), \quad (A2)$$

we can write

$$N^{n}(\vec{U})\Psi(\vec{U}) \propto \exp\left(-\frac{\alpha}{2}(\vec{u}-\vec{u}_{0})^{2}-\frac{\gamma}{2}\vec{v}^{2}\right), \quad (A3)$$

with

$$\alpha = 1 + \alpha_n = 1 + \frac{n}{\tanh(2\tau_1 + \tau_2)} - \frac{n-1}{\tanh\tau_2}, \qquad (A4)$$

$$\beta_n = \frac{n}{\sinh(2\tau_1 + \tau_2)} - \frac{n-1}{\sinh\tau_2},\tag{A5}$$

$$\vec{u}_0 = \frac{\beta_n}{\alpha} \vec{v}, \qquad (A6)$$

$$\gamma = \alpha - \beta_n \,. \tag{A7}$$

Since $\alpha_n > \beta_n$, it is clear that the only condition that must hold for M_n to be finite is that α be positive. This easily leads to

$$\tanh(2\tau_1 + \tau_2) < \frac{n \tanh \tau_2}{n - 1 - \tanh \tau_2},\tag{A8}$$

- A. Aspuru-Guzik and W.A. Lester, Jr., e-print, arXiv:cond-mat/ 0204486.
- [2] J.B. Anderson, J. Chem. Phys. 112, 9699 (2000).
- [3] S.C. Pieper and R.B. Wiringa, Annu. Rev. Nucl. Part. Sci. 51, 53 (2001).
- [4] M. Foulkes, L. Mitas, R. Needs, and G. Rajagopal, Rev. Mod. Phys. 73, 33 (2001).
- [5] D.M. Ceperley and B.J. Adler, J. Chem. Phys. 81, 5833 (1984).
- [6] M.H. Kalos and F. Pederiva, Phys. Rev. Lett. 85, 3547 (2000);
 Physica A 279, 236 (2000).
- [7] P.A. Whitlock, D.M. Ceperley, G.V. Chester, and M.H. Kalos, Phys. Rev. B 19, 5598 (1979).
- [8] P.J. Reynolds, R.N. Barnett, B.L. Hammond, R.M. Grimes, and W.A. Lester, Jr., Int. J. Quantum Chem. 29, 586 (1986).
- [9] K.S. Liu, M.H. Kalos, and G.V. Chester, Phys. Rev. A 10, 303 (1974).
- [10] P.J. Reynolds, R.N. Barnett, B.L. Hammond, and W.A. Lester, Jr., J. Stat. Phys. 43, 1017 (1986).
- [11] R.N. Barnett, P.J. Reynolds, and W.A. Lester, Jr., J. Comput. Phys. **96**, 258 (1991).
- [12] K.J. Runge, Phys. Rev. B 45, 7229 (1992).
- [13] J. Casulleras and J. Boronat, Phys. Rev. B 52, 3654 (1995).

the first condition mentioned in the text. Since tanh(x) < 1, this condition is simple if τ_2 is large; that is, if

$$\tau_2 > \frac{1}{2} \ln n, \tag{A9}$$

the condition is satisfied for every τ_1 .

Equation (A8) is not easy to interpret analytically in the short-time limit. We can obtain an useful form by neglecting the tanh τ_2 in its denominator:

$$\frac{\tanh(2\,\tau_1+\tau_2)}{\tanh\tau_2} \leqslant \frac{n}{n-1}.\tag{A10}$$

This is a more restrictive condition; as a matter of fact, this condition keeps the integral

$$\int d\vec{u} d\vec{v} \frac{G_{1,1}^{n}(\vec{u},\vec{v},2\tau_{1}+\tau_{2})}{G_{1,1}^{n-1}(\vec{u},\vec{v},\tau_{2})}$$

finite without considering the effect of the ground-state wave functions as in Eq. (A1). In the short-time limit, Eq. (A10) becomes

$$\tau_2 \geq 2(n-1)\tau_1$$

which is the second equation mentioned in the text.

- [14] P. Langfelder, S.M. Rothstein, and J. Vrbik, J. Chem. Phys. 107, 8525 (1997); M. Hornik, M. Snajdr, and S.M. Rothstein, *ibid.* 113, 3496 (2000).
- [15] A. Sarsa, K.E. Schmidt, and W.R. Magro, J. Chem. Phys. 113, 1366 (2000).
- [16] B.H. Wells, Chem. Phys. Lett. 115, 89 (1985).
- [17] C.A. Traynor and J.B. Anderson, Chem. Phys. Lett. 147, 389 (1988).
- [18] S. Zhang and M.H. Kalos, J. Stat. Phys. 70, 515 (1993).
- [19] R.P. Feynman, *Statistical Mechanics: A Set of Lectures* (Addison-Wesley, Reading, MA, 1992).
- [20] S.T. Epstein, A.C. Hurley, R.E. Wyatt, and R.G. Parr, J. Chem. Phys. 47, 1275 (1967).
- [21] B.L. Hammond, W.A. Lester, Jr., and P.J. Reynolds, *Monte Carlo Methods in Ab Initio Quantum Chemistry* (World Scientific, Singapore, 1994).
- [22] J.B. Anderson, J. Chem. Phys. 63, 1499 (1975); 65, 4121 (1976).
- [23] F. Arias de Saavedra, I. Porras, E. Buendía, and F.J. Gálvez, J. Phys. B 28, 3123 (1995).
- [24] T. Kato, Commun. Pure Appl. Math. 10, 51 (1957).