Shifted-contour auxiliary field Monte Carlo for ab initio electronic structure: Straddling the sign problem

Roi Baer, Martin Head-Gordon, and Daniel Neuhauser

Articles you may be interested in

A field-theoretical approach to simulation in the classical canonical and grand canonical ensemble
The Journal of Chemical Physics **117**, 3027 (2002); 10.1063/1.1488587

Excited state calculations using phaseless auxiliary-field quantum Monte Carlo: Potential energy curves of low-lying C_2 singlet states
The Journal of Chemical Physics **130**, 094107 (2009); 10.1063/1.3077920

Shifted-contour auxiliary field Monte Carlo for *ab initio* electronic structure: Straddling the sign problem

Roi Baer and Martin Head-Gordon
Department of Chemistry, University of California and the Chemical Sciences Division,
Lawrence Berkeley National Laboratory, Berkeley, California 94720

Daniel Neuhauser[a)]
Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095-1569

The auxiliary field Monte Carlo (AFMC) technique has advantages over other *ab initio* quantum Monte Carlo methods for fermions, as it does not seem to require approximations for alleviating the sign problem and is directly applicable to excited states. Yet, the method is severely limited by a numerical instability, a numerical sign problem, prohibiting application to realistic electronic structure systems. Recently, the *shifted contour auxiliary field method* (SC-AFMC) was proposed for overcoming this instability. Here we develop a theory for the AFMC stabilization, explaining the success of SC-AFMC. We show that the auxiliary fields can be shifted into the complex plane in a manner that considerably stabilizes the Monte Carlo integration using the exact one-electron density. Practical stabilization can be achieved when an approximate Hartree–Fock density is used, proposing that an overwhelming part of the sign problem is removed by taking proper account of the Fermion mean-field contribution. The theory is demonstrated by application to H_2. © *1998 American Institute of Physics.* [S0021-9606(98)00539-X]

I. INTRODUCTION

Quantum Monte Carlo (QMC) methods offer a nonperturbative means for systematically improving the quality of electronic structure calculations, taking into account many-body correlations, while scaling gently with number of electrons. QMC methods for electronic structure calculations are the variational (VMC),[1–3] Green's function (GFMC),[4–7] diffusion (DMC),[8,9] path integral (PIMC),[10,11] Auxiliary Fields (AFMC)[12] and Constrained Path (CPMC).[13–15] Although these approaches differ in representation of the wave functions and projection operators, a universal feature is the occurrence of a "fermion sign problem," limiting the accuracy or stability of the methods for electronic structure. In PIMC, for example, it is necessary to change the sign of all odd permutations, and when the temperature is lowered to the Fermionic temperature a catastrophic loss of accuracy occurs. In DMC, GFMC, and CPMC the sign problem is associated with the need to represent a sign change of the exact electronic ground state when random walkers cross the wave function nodes. In this context, the fixed node method[8] has been a very successful approximation, occasionally providing tight upper bounds to the full correlation energy. A special version of the fixed node approximation, the restricted path MC, has allowed a stabilization of the PIMC.[10]

VMC, DMC, and GFMC are by now well-established methods and have demonstrated their power in many applications.[16–19] There is, however, more to be desired. One of the goals of Monte Carlo methods is to overcome the exponential scaling of full CI methods with respect to

accuracy.[19] The fixed node approximation fails in this aspect because the error bar can be decreased only by generating a better node structure, which, at present, is an exponentially hard computing problem.[19] Furthermore, DMC methods are especially useful only for energy calculations. Other observables are biased by the importance function.[19] Last, despite some progress made in developing methods for excited states,[20,17] these have not yet showed convincing generality, precision, and robustness for realistic chemical systems.[19] Thus, there is a real need for additional QMC approaches that may cope better with the sign problem and allow for efficient calculation of a variety of observables.

In this paper we address what seems a likely candidate for the task: the Auxiliary Fields Monte Carlo (AFMC) method. AFMC is especially attractive because the sign problem seems to be "analytically taken care of," without uncontrolled approximations such as trial wave functions. Furthermore, the formalism deals with imaginary as well as real time evolution, so excited states and other dynamical observables are exactly obtainable. The method uses Monte Carlo (MC) techniques for evaluating the functional integrals appearing in the Hubbard–Stratonovich[21,22] representation of the evolution operator. AFMC is applicable to fermions (and bosons) in an external potential, interacting among themselves via two body forces.

For fermions, the AFMC method has been used extensively to study electron correlation in the Hubbard model.[23–28] It was found that some types of electron–electron interactions on Hubbard lattices yield semi positive-definite integrands and were amenable to calculation via the AFMC method.[25,28] However, when applied to atoms and molecules, with a realistic electron–electron Coulomb repul-

[a)]Electronic mail: dxn@chem.ucla.edu

sion, it was observed that the Monte Carlo integration was next to impossible, even for the helium atom[29] (as well as most repulsive Hubbard models). Even the most basic quantum chemical application, that of calculating the bond energy of the H_2 ground state,[30] seems to have encountered great difficulties, since no attempt to converge the correlation energy with respect to lowering the temperature was attempted. Thus, although analytically ''taken care of,'' the sign problem is not eliminated in AFMC but simply transformed. It becomes a numerical sign problem: when an attempt to lower the temperature in the AFMC integrand is made, the statistical fluctuations become exponentially large.[28]

The sign problem in AFMC was also analyzed by Fahy and Hamann,[31] where it was shown that the AFMC method is isomorphic to a diffusion–decay–drift process on the manifold of normalized Slater determinants. In this picture the origin of the sign problem is traced back to the natural inclination of a diffusing liquid to eventually assume a smooth nodeless form, while the Fermionic ground state contains a complicated oscillatory form due to the nodes caused by the Pauli exclusion principle. The authors also find that the ground state energy of the noninteracting Hamiltonian plays a dominant role in the diffusion drift–decay determining potentials. This interesting approach has led to the formulation of a brand of new AFMC methods, the positive projection (PPMC)[13] and constrained path Monte Carlo (CPMC),[14,15] where the AFMC instability is eliminated by a fixed-node type of approximation in the determinantal manifold.

In this paper we attempt to contribute to efforts invested in coping with the sign problem in QMC calculations. We show how it is possible to effectively stabilize the large-amplitude oscillations in the AFMC integrand. Unlike the Fahy–Hamann formulation, the approach here conforms to the spirit of the original Hubbard–Stratonovich (HS) transformation and there is no need for a geometric construction of a determinantal manifold. Thus, the AFMC sign problem and the question of the role of the mean field is examined from a completely different perspective yielding new results and insights.

The SC-AFMC ideas were originally proposed in Ref. 32 for stabilizing the AFMC integrand. Here we give a more detailed treatment. We show in Sec. III that one can considerably stabilize the AFMC integrand by using an exact time-dependent matrix element of the complete one-electron density matrix. Of course, in practice, the exact one-electron density is not known, so an approximation is used. Depending on the approximation, this may allow a sufficiently accurate result to be obtained.

The theory is applied to the H_2 molecule with orbitals represented on a grid. Using the Hartree–Fock density we were able to extract 99% of the correlation energy describable by the grid representation used. A previous application of AFMC[30] did not attempt to collect the complete correlation energy. We show that such an attempt would have been doomed to failure unless special stabilization techniques, such as SC-AFMC, are used.

Structure of this paper: the AFMC formalism is briefly presented in Sec. II, and the theory for stabilizing AFMC

calculations is derived in Sec. III. Numerical issues are discussed in Sec. IV. The application of the theory to H_2 is studied in Sec. V. Finally, conclusions are discussed in Sec. VI.

II. OVERVIEW OF AFMC FORMALISM IN R SPACE

In AFMC, the central physical quantity for electronic structure is the imaginary time evolution operator, or partition function operator for the electrons:

$$U = \exp\{-\beta \hat{H}\}, \tag{2.1}$$

where β is an imaginary time or an inverse temperature. The full many-electron Hamiltonian is

$$\hat{H} = \sum_s \int \hat{\Psi}_s^+(x)\left(\frac{-1}{2m_e}\nabla^2 + w(x)\right)\hat{\Psi}_s(x)d^3x$$

$$+ \frac{e^2}{2}\sum_{s,s'}\int\int \hat{\Psi}_s^+(x)\hat{\Psi}_{s'}^+(y)\frac{1}{|x-y|}\hat{\Psi}_{s'}(y)\hat{\Psi}_s(x)d^3x\,d^3y, \tag{2.2}$$

where m_e and e are, respectively, the electron mass and charge, $w(x)$ is the electrostatic potential of the nucleon–electron interaction, and $\hat{\Psi}_s^+(x)$ creates an electron of spin s at x while $\hat{\Psi}_s(y)$ annihilates a similar electron at y. Both x and y are coordinates of an electron in its three-dimensional position space.

The trace of the operator in Eq. (2.1) gives the partition function:

$$Z(\beta) = tr\{U(\beta)\} = \sum_k \langle \Phi_k|U(\beta)|\Phi_k\rangle, \tag{2.3}$$

from which the free energy, $F(\beta) = -\partial \ln Z/\partial\beta$ and the ground state energy, $E_{gs} = \lim_{\beta\to\infty} F(\beta)$, can be extracted. The sum in Eq. (2.3) is over any complete set of linearly independent N-particle wave functions, such as a complete set of determinants. If only the ground state energy is needed, one can use as few as one term in Eq. (2.3).

The Hamiltonian of Eq. (2.2) can be written (up to some constants) in the following compact form, needed for the AFMC formalism:

$$\hat{H} = K^T\hat{\rho} + \tfrac{1}{2}\hat{\rho}^T V\hat{\rho}. \tag{2.4}$$

Here $\hat{\rho}(x,y) = \Sigma_s \hat{\Psi}_s^+(x)\hat{\Psi}_s(y)$ is an element of the one-electron density matrix describing quantum correlations of the electron density at points x and y. The pair of position vectors (x,y) may be viewed as a single index and the ''inner product'' of a vector so indexed is compactly denoted by

$$A^T B \equiv \int dx\,dy\,A(x,y)B(x,y). \tag{2.5}$$

The vector $K(x,y) = \delta(x-y)[-(1/2m_e)\nabla_x^2 + w(x)]$ in Eq. (2.4) includes all the one-body operators (nuclear attraction w and the kinetic energy), and V is the matrix describing electron–electron Coulomb repulsion:

$$V(x,y;x',y') = e^2\frac{\delta(x-y)\delta(x'-y')}{|x-x'|}. \tag{2.6}$$

The form of Eq. (2.4) for the Hamiltonian can be attained for any one-particle representation,[12] so the stability discussion encompasses all forms of AFMC implementations.

Casting of the Hamiltonian in terms of density operators enables us to use the HS transformation:[21,22,33,34]

$$e^{-(1/2)A^T \mathbf{W} A} = \left[\frac{1}{2\pi}\right]^{N/2} \sqrt{\det(\mathbf{W})}$$
$$\times \int e^{-(1/2)\sigma^T \mathbf{W}\sigma} e^{-i\sigma^T \mathbf{W}A} d^N\sigma, \qquad (2.7)$$

where $\mathbf{W}$ is a positive definite matrix. This equality is easily proved using usual Gaussian integration, when the vector A is composed of N numbers. When we insert, instead of the vector A the vector ρ of N noncommuting one-body operators, the equality can be preserved in the limit of a small exponent:

$$e^{-(1/2)\rho^T \mathbf{V}\rho \,\Delta t} = \left[\frac{\Delta t}{2\pi}\right]^{N/2} \sqrt{\det(V)}$$
$$\times \int e^{-(1/2)\sigma^T V\sigma \,\Delta t} e^{-i\sigma^T \mathbf{V}\,\rho\Delta t} d^N\sigma, \quad (2.8)$$

where Δt is infinitesimal. Thus, a single imaginary time slice of an evolution operator for a Hamiltonian of the two-body interaction $\frac{1}{2}\rho^T V\rho$ is written as a sum of real time evolution operators that involve only one-particle interactions of the form $h_\sigma = \sigma^T V\rho$ and include an auxiliary field $V\sigma$. The integral is over all possible values of the auxiliary densities σ and is weighted by a Gaussian, $\exp(-\frac{1}{2}\sigma^T V\sigma\,\Delta t)$.

For a finite time β we must divide into time slices and every slice involves a separate auxiliary field integral. Taking the number of one-body operators to a continuous limit, and adding a one-body interaction, the matrix element of the full evolution operator can be represented as a multidimensional functional integral over all possible time-dependent auxiliary fields $\sigma_\tau(x,y)$:

$$\langle\Phi|e^{-\beta(K^T\rho + (1/2)\rho^T \mathbf{V}\rho)}|\Phi\rangle = \int W[\sigma_\tau] U^{(1)}_{\sigma_\tau}(\beta) D\{\sigma_\tau\}$$
$$\equiv \langle U^{(1)}_{\sigma_\tau}(\beta)\rangle_W. \qquad (2.9)$$

In most applications the N electron state is taken as a determinant,[12] $\Phi|\psi_1\cdots\psi_N|$, of N spin orbitals ψ_n, and the functional integral is calculated using Monte Carlo techniques.

Equation (2.9) is a tremendous simplification because the term $U^{(1)}_{\sigma_\tau}(\beta)$ is a matrix element of the evolution operator of a one-particle time-dependent Hamiltonian containing the auxiliary fields $V\sigma_\tau(x,y)$:

$$U^{(1)}_{\sigma_\tau}(\beta) = \left\langle \Phi\left|\hat{T}\,\exp\left\{-\int_0^\beta (K+iV\sigma_\tau)^T\hat{\rho}\,d\tau\right\}\right|\Phi\right\rangle \qquad (2.10)$$

($\hat{T}$ is the time ordering operator). Thus, when the time-dependent evolution operator operates on a determinant Φ, the result is again a determinant $\Phi(\beta) = |\psi_1(\beta)\cdots\psi_N(\beta)|$, composed of the separately propagated orbitals:

$$\hat{T}\,\exp\left\{-\int_0^\beta (K+iV\sigma_\tau)^T\hat{\rho}\,d\tau\right\}\psi_n = \psi_n(\beta). \qquad (2.11)$$

The fluctuating fields, $iV\sigma_\tau(x,y)$ in Eq. (2.9), are responsible for building the correct electron–electron interaction into the evolution operator matrix elements.

Numerical evaluation of the HS functional integral of Eq. (2.9) uses Monte Carlo methods. Because quantum evolution is Markovian, the auxiliary fields of different time slices are statistically independent in the Monte Carlo approach. In each time slice $t \to t+\Delta\tau$ the auxiliary fields are sampled according to a Gaussian distribution W:

$$W[\sigma_\tau] = e^{-(1/2)\sigma_\tau^T V\sigma_\tau \,\Delta\tau}, \qquad (2.12)$$

which is normalized through the differentials $D\{\sigma_\tau\}$:

$$\int W[\sigma_\tau] D\{\sigma_\tau\} = 1. \qquad (2.13)$$

Once a field is sampled, the right-hand side determinantal wave function $|\Phi(t)\rangle$ is propagated by an additional time step $\Delta\tau$ [see Eq. (2.11)]. This process continues until one reaches the determinant $|\Phi(\beta)\rangle$. Then the overlap between determinants is calculated as $\langle\Phi|\Phi(\beta)\rangle$. This last quantity is repeatedly summed over, statistically converging to the diagonal matrix element of the many-body evolution operator.

This elegant formalism has advantages over present MC approaches in that it assumes no approximate form for the electronic wave function or for its nodes as do VMC and fixed node DMC. The use of determinants is exact, not an approximation. However, these hypothetical merits are not truly realized since the sign problem has not disappeared, it merely changed form: now it is manifested as a numerical instability. We address this issue and its resolution in the following section.

III. STABILIZING THE FUNCTIONAL INTEGRANDS: SC-AFMC

The excited states can be damped relative to the ground state in two general ways. One way is to use real time evolution to cause the excited states to oscillate rapidly and then sum over these oscillations, canceling their effect. The other way is to dampen high-energy terms in the wave functions by an imaginary time evolution (like heat transfer or diffusion). The drawback of the HS transformation of Eq. (2.9) is that it damps the high-energy components in the first way mentioned above. Thus as seen in Eq. (2.10), it changes imaginary time evolution into a summation over a real time evolution.

Our goal is then to modify the HS transformation in such a way as to prevent the formation of large-amplitude oscillations in the integrand. In order to achieve this goal we must figure out a way to accurately compensate, in terms of imaginary-time damping, a reduction of the real time oscillations. This can be done by invoking the well-known complex plane path invariance of line integrals. It is straightforward to prove that the functional integral of Eq. (2.9) is invariant to an imaginary contour shift $i\alpha_\tau(x,y)$, an arbitrary real field that may also depend on time:

$$\langle U^{(1)}_{\sigma_\tau}(\beta)\rangle_W = e^{1/2\int_0^\beta \alpha_\tau^T \mathbf{V}\alpha_\tau \, d\tau}\left\langle e^{i\int_0^\beta \sigma_\tau^T \mathbf{V}\alpha_\tau \, d\tau} U^{(1)}_{\sigma_\tau - i\alpha}(\beta)\right\rangle_W. \tag{3.1}$$

A close examination of Eq. (3.1) shows that the contour shift $i\alpha_\tau$ appears in three places: in the one-body evolution operator $U^{(1)}$, as part of an oscillating phase factor, and finally as an exponential amplification factor. Notice that within the evolution operator the contour shift induces damping while in the phase factor it induces an oscillation; this allows us to control the balance between oscillations and damping. The exponential amplification will be ignored in subsequent discussions because in all expectation value expressions it cancels in the numerator and denominator.

The question is now: how should the imaginary fields $i\alpha_\tau(x,y)$ be chosen to mitigate the oscillations of the integrand? In order to answer that, let us write a matrix element of the evolution operator in the following form:

$$\langle \Psi(0)|U(\beta)|\Phi(0)\rangle \propto \int D\sigma_\tau \, W[\sigma_\tau]e^{-iS_\tau \Delta t}. \tag{3.2}$$

The functional integral now is over the field of a single time slice $\tau \to \tau + \Delta t$. Each time slice Δt is separately stabilized and the complex action S_τ is defined as

$$S_\tau = i\frac{\ln\langle \Psi(\beta-\tau)|e^{-(K+iV(\sigma_\tau+i\alpha_\tau))^T\hat{\rho}\Delta t}|\Phi(\tau-\Delta t)\rangle}{\Delta t}$$
$$- \sigma_\tau^T V \alpha_\tau. \tag{3.3}$$

Our plan is to choose the imaginary field $i\alpha_\tau$ so as to have S_τ insensitive to first-order changes in the auxiliary field σ_τ. The sensitivity of S_τ to σ_τ is the following functional derivative:

$$\frac{\delta S_\tau}{\delta \sigma_\tau} = iV\left(\alpha_\tau - \frac{\langle \Psi(\beta-\tau)|\hat{\rho}|\Phi(\tau)\rangle}{\langle \Psi(\beta-\tau)|\Phi(\tau)\rangle}\right), \tag{3.4}$$

and thus a sufficient condition for S_τ to be stationary with respect to σ_τ is

$$\alpha_\tau(x,y) = \frac{\langle \Psi(\beta-\tau)|\hat{\rho}(x,y)|\Phi(\tau)\rangle}{\langle \Psi|\Phi(\beta)\rangle}. \tag{3.5}$$

The field shift α_τ is then equal to a matrix element of the exact one-electron density matrix normalized by the matrix element of the evolution operator. Notice that when β, τ, and $\beta - \tau$ are very large, we have

$$\alpha_\tau(x,y) = \langle \Theta_{gs}|\hat{\rho}(x,y)|\Theta_{gs}\rangle \quad (\beta,\tau,\beta-\tau \to \infty), \tag{3.6}$$

where Θ_{gs} is the exact normalized ground state wave function. Thus, in these limits α_τ is the exact ground state one-electron density $\langle\hat{\rho}\rangle_{gs}$. When β is indefinitely large, α_τ becomes equal to the *exact* ground state density at all times.

Note that due to the special form of V for the Coulomb interaction [see Eq. (2.6], only the diagonal elements of α_τ [i.e., $\alpha_\tau(x) = \alpha_\tau(x,x)$] affect the functional integrals. The sufficient condition of Eq. (3.5) is then simplified to a condition pertaining only to the electronic density:

$$\alpha_\tau(x) = \frac{\langle \Psi(\beta-\tau)|\hat{\rho}(x)|\Phi(\tau)\rangle}{\langle \Psi|\Phi(\beta)\rangle}. \tag{3.7}$$

When the AFMC calculation uses a high-β parameter, for computation of ground state properties, a simple approximation to the ground-state one-electron density may be sufficient for achieving practical stability in the Monte Carlo integration. In many systems the density is well approximated by Hartree–Fock theory. This is the idea proposed in Ref. 32, where a ten-electron neon atom was treated using the HF density to stabilize the functional integral. In nondegenerate systems the closed-shell restricted Hartree-Fock (RHF) or, when needed, the unrestricted Hartree–Fock (UHF) may be used to obtain good approximations to the density. It is perhaps possible also to use density functional theory (DFT) with the local spin-density approximation (LSD) for this purpose.

IV. NUMERICAL METHOD

In Sec. V we demonstrate the theory using the hydrogen molecule as an example. In this section we discuss numerical representation and implementation issues.

A. Grid representation

Consider the molecule in a cubic cell of dimensions L^3. Our approximations in representing the system become exact in the large L limit. Several results from the plane waves method[35] are used, which enhance numerical accuracy and stability even for more moderate values L. We assume that the cell is surrounded with identical systems which together form a superlattice. The electronic state in all cells is identical ($k=0$ approximation). In order to eliminate the divergences in the electron–electron, electron–nucleon and nucleon–nucleon potentials, we zero the q=0 Fourier component of all these interactions.[35] Thus, the R-space potential for the electron–nucleon interaction (taking into account all nucleons in the superlattice) is

$$v_{\text{ext}}(r) = \frac{1}{\sqrt{L^3}}\sum_{q\neq 0} v_q e^{-irq}, \tag{4.1}$$

where e^{iqr} are standing plane waves inside the cell, and

$$v_q = -\frac{1}{\sqrt{L^3}}\sum_m \frac{4\pi e^2 Z_m e^{iR_m q}}{q^2}, \quad q\neq 0. \tag{4.2}$$

Here m indexes the nuclei of the system cell, each at location R_m. The two-electron interaction, is also taken as

$$V_{ee}(r) = \frac{1}{\sqrt{L^3}}\sum_{q\neq 0} V_q e^{-irq}, \tag{4.3}$$

where

$$V_q = \frac{1}{\sqrt{L^3}}\frac{4\pi e^2}{q^2}, \quad q\neq 0. \tag{4.4}$$

The nuclear–nuclear interaction is calculated via an Ewald summation excluding the q=0 component. Since we use a very large cell, it is consistent to assume that the electronic
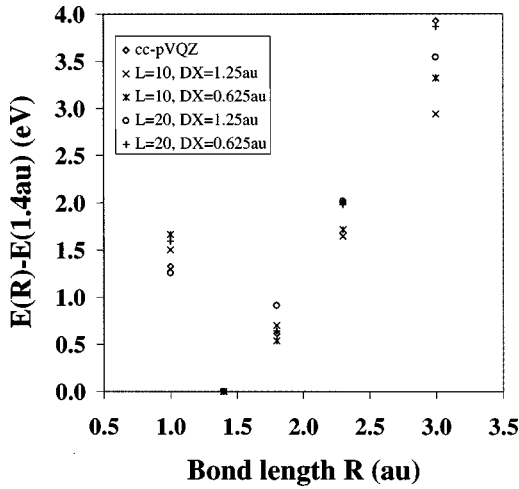
FIG. 1. Restricted Hartree–Fock total (electronic and nuclear repulsion) energy above minimum of H_2 using a basis set (cc-pVQZ) and 2×2 combinations of cell lengths (L) and grid spacing (Δx).

Coulomb potential operates on localized functions within the cell and there is no need to use an Ewald sum for interactions involving electrons. The R-space functions are represented on a grid,[36] which naturally introduces a cutoff in the q expansions via the finite grid sampling Δx.

The faithfulness of this representation can be studied by comparing its performance to that of a high-quality Gaussian basis at the Hartree–Fock level. For the Gaussian basis calculation we used the program Q-CHEM[37] with the cc-pVQZ basis of Ref. 38. In Fig. 1 we present the Hartree–Fock energies of the basis set calculation and four cases of grid calculations. For the latter we used two values of cell sizes L and 2 values of grid spacing Δx. It is seen that for a broad range of bond lengths, the plane wave results indeed converge to the basis set calculations as the cell size is increased and grid spacing is decreased. However, the convergence is rather slow. For larger than equilibrium bond lengths the convergence is slow with respect to cell size, while for smaller than equilibrium bond length it is slow with respect to grid spacing. The former phenomenon is typical for Hartree–Fock calculations using plane waves and grid methods.[39] Several refinements are possible to reduce these errors,[39–41] however, these were not implemented in the present work.

It has been established, at least for the homogeneous electron gas, that finite size errors in fully correlated energies are similar to those of the Hartree–Fock calculation.[40] This motivates our assumption that the correlation energy is less sensitive to finite size errors and can therefore be studied qualitatively using a moderate-sized cell ($L = 10$ a.u.). This is in accordance to our goal in this paper, to focus on the *statistical convergence* of the AFMC calculation. Correspondingly, we shall also use the larger grid spacing value of $\Delta x = 1.25$ a.u.

B. The SC-AFMC algorithm

The SC-AFMC computation starts by first calculating the SCF Hartree–Fock density. In the present application we use a closed shell formalism and solve for the lowest energy occupied molecular orbitals $\psi_1 \cdots \psi_N$, where $2N$ is the number of electrons (in our case $N = 1$). The density matrix is then

$$\alpha(x) = 2 \sum_{n=1}^{N} |\psi_n(x)|^2. \tag{4.5}$$

This electron density is the stabilizing shift, and the derived stabilizing potential is

$$\Xi(x) = (V\alpha)(x) = \int \frac{\alpha(y)}{|x-y|} dy. \tag{4.6}$$

In practice, fast Fourier transform (FFT) is used to perform this spatial integral, using the Fourier representation of the two-electron potential in Eq. (4.4). Using closed shell Hartree Fock theory limits our ability to accurately describe the mean field density at large bond lengths, when the ground state consists of two open shells. This density is not expected to efficiently stabilize the AFMC calculation. We return to this point in Sec. VI.

Next, the computationally intensive part of the calculation begins. In this part the following expression for the ground state energy is calculated:

$$E_{\text{gs}} = \frac{\langle \Phi | H U(\beta) | \Phi \rangle}{\langle \Phi | U(\beta) | \Phi \rangle}, \tag{4.7}$$

where H is the full many-electron Hamiltonian and $\Phi = |\psi_1 \bar{\psi}_1 \cdots \psi_N \bar{\psi}_N|$ is the closed shell determinant composed from the N occupied orbitals. We mention in passing that since the spatial part of ψ_n and $\bar{\psi}_n$ are identical, only one of the orbitals actually needs to be operated on by the AFMC one-body (spin independent) evolution operators.

The imaginary time β is divided to M equal segments, $\Delta \tau = \beta / M$. In the first segment a random auxiliary field is sampled from a Gaussian distribution:

$$W\{\sigma(x)\} = A e^{-(1/2)\sigma^T V \sigma \, \Delta t}, \tag{4.8}$$

where A is a normalizing factor. The sampling is efficiently achieved by decoupling the different fields using a fast Fourier transform. The sampled field σ is used to construct the auxiliary interaction potential:

$$W(x) = (V\sigma)(x) = \int \frac{\sigma(y)}{|x-y|} dy. \tag{4.9}$$

Notice that owing to the special diagonal form of the V matrix in Eq. (2.6), the auxiliary field depends only on a single continuous index (instead of two). This is contrary to a basis set AFMC and is a consequence of the special form the two body-interaction has in the R-space representation. This is important for a favorable scaling of numerical work with a number of electrons.

Once the field is sampled, the resulting one-electron evolution operator,

$$\exp\left[-\left(-\frac{1}{2m_e} \nabla^2 + v_{\text{ext}}(x) + i[W(x) - i\Xi(x)] \right) \Delta \tau \right],$$

is applied, using Chebyshev propagation methods,[42–44] to each of the N orbitals in the determinant to obtain the determinant (multiplied by the stabilizing phase factor):

$$\Phi_{\Delta\tau} = e^{i\sigma^T W\sigma \, \Delta\tau} \left| \psi_1(\Delta\tau)\bar{\psi}_1(\Delta\tau)\cdots\Psi_N(\Delta\tau)\bar{\psi}_N(\Delta\tau) \right|. \tag{4.10}$$

This procedure is repeated for every one of the M time segments. The final result Φ_β is a propagated determinant under a particular choice of auxiliary fields, multiplied by the stabilizing phase factors. The overlap $S = \langle\Phi|\Phi_\beta\rangle$ and the energy $E = \langle\Phi|H|\Phi_\beta\rangle$ are then calculated, thus concluding one Monte Carlo iteration. Repeating this procedure and summing over all sampled contributions E and S finally gives the estimated ground state energy:

$$E_{gs} = \frac{\Sigma E}{\Sigma S}. \tag{4.11}$$

V. RESULTS

A. SC-AFMC correlation energy of H_2

At the equilibrium bond length of $R = 1.4$ a.u., we have tested several values of the time interval $\Delta\tau$ and found that it is sufficient to use a value of $\Delta\tau = 0.1$ a.u. The dependence of the calculated ground state energy $E(\beta)$ can be estimated by the following two-level model at large β values:

$$\frac{E(\beta) - E_{gs}}{\Delta E} \approx \left| \frac{\langle\Psi_1|\Phi\rangle}{\langle\Psi_0|\Phi\rangle} \right|^2 e^{-\Delta E \, \beta} = q^2 e^{-\Delta E \, \beta}, \tag{5.1}$$

where E_{gs} is true the ground state energy, ΔE the molecular excitation energy, and Ψ_0 and Ψ_1 are, respectively, the ground and first-excited wave functions. Φ is the initial AFMC determinant, taken in these calculations as the Hartree–Fock ground state. From Eq. (5.1) we see that the ground state energy is recovered at a rate determined by the excitation energy ΔE. Using SC-AFMC we have calculated the energy $E(\beta)$ for four values of β. A least squares procedure allowed us to obtain the parameters E_{gs}, ΔE, and q^2 of Eq. (5.1):

$$E_{gs} - E_{HF} = -0.0365(4) \text{ a.u.} = -0.99(2) \text{ eV},$$

$$\Delta E = 0.913 \text{ a.u.}, \tag{5.2}$$

$$q^2 = 0.014(1) \text{ a.u.}$$

Numbers in parentheses are the statistical errors in the last shown digit. For comparison purposes, the correlation energy for the cc-pVQZ basis is -1.1 eV. Thus, the 512 point grid can account for 90% of the full correlation energy.

The statistical and temporal convergence properties are shown in Fig. 2. The results reach chemical accuracy at about $\beta = 8$ a.u. All the statistical errors are on the order of 0.02 eV or less and grow moderately with the length of the propagation. The correlation energy at $\beta = 1$ a.u. was calculated to be 0.0313 a.u. and is different from the value of 0.035 a.u. reported by Silverstrelli et al.[30] This could be caused by differences in implementation details of the two calculations (grid and operator definitions, and especially the treatment of the boundary conditions and finite size effects).
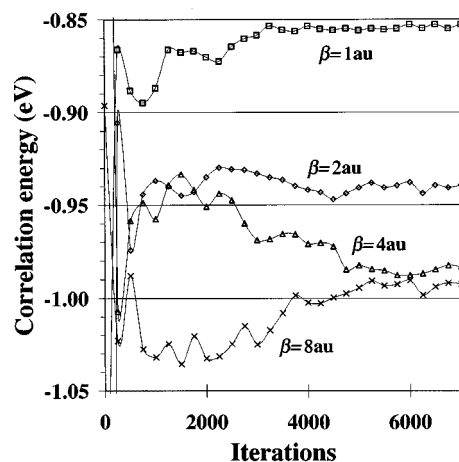

FIG. 2. SC-AFMC results of the correlation energy of H_2 at a bond length of 1.4 a.u. Only results for iterations number 250, 500, 750, etc. are shown and the connecting lines are a guide to the eye.

It is interesting to compare the statistical covergence of the correlation energies at the equilibrium bond length $R = 1.4$ a.u. (results shown in Fig. 2) with those of the almost dissociated molecule when the bond length is $R = 3.0$ a.u. In the latter case the RHF density is not a good approximation of the true electronic density, and the integrand stabilization should be damaged. This is indeed seen in Fig. 3, which exhibits inferior convergence when compared with Fig. 2.

B. The sign problem

The result of an AFMC calculation (without the Shifted Contour) with β parameter equal to 1, 2, and 4 a.u. are shown in Fig. 4. For the $\beta = 1$ a.u. case, the AFMC shows a reasonable result, although the statistical errors are much larger than those of the SC-AFMC. However, only 80% of the correlation energy can be retrieved by a $\beta = 1$ a.u. calculation (see Fig. 2). It is when one attempts to increase β that the sign problem really becomes apparent. This is seen in Fig. 4 as an extremely rapid loss of stability with increasing β. The truly superior properties of SC-AFMC are now evident by comparing Fig. 4 and Fig. 2 for $\beta = 2$ a.u.; it is seen
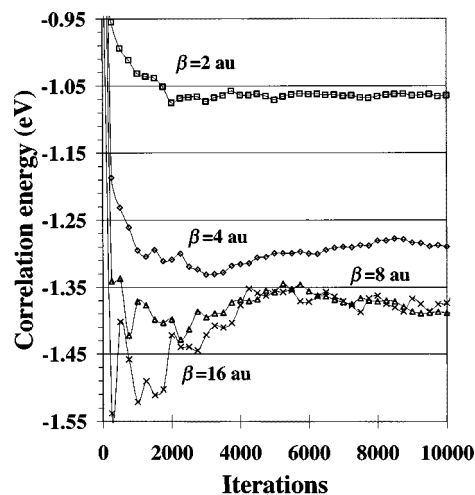

FIG. 3. The same as Fig. 2, for a bond length of 3.0 a.u.
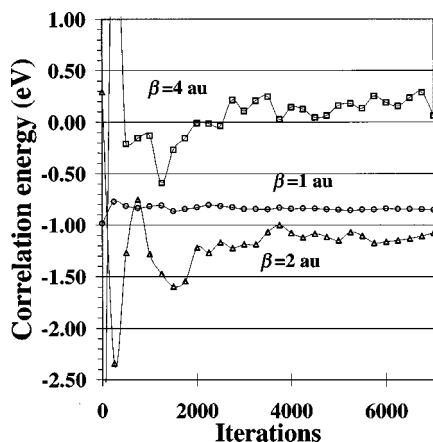
FIG. 4. The same as Fig. 2, but using AFMC with no contour shift. Statistical convergence degrades rapidly as β increases. The $\beta = 4$ a.u. results are hopelessly far from convergence (show positive correlation energy).

FIG. 5. The sign Σ_α for H_2 as a function of β. Triangles: no contour shift; filled dots and squares: contour shift determined by close-shell restricted Hartree–Fock density.

that the AFMC statistical convergence is greatly degraded while that of SC-AFMC only slightly. At $\beta = 4$ a.u. the AFMC calculation becomes totally unstable, predicting positive correlation energy. This is in contrast to the SC-AFMC calculations that remain stable (even up to $\beta = 16$ a.u., the largest value we checked).

Another way to see this remarkable stabilization is by examining the quantity called ''the sign,'' defined as:[28]

$$\Sigma_0(\beta) = \frac{\langle U_\sigma^{(1)}(\beta) \rangle_W}{\langle |U_\sigma^{(1)}(\beta)| \rangle_W} = \frac{\langle \Phi | U(\beta) | \Phi \rangle}{\langle |U_{\sigma-i\alpha}^{(1)}(\beta)| \rangle_W}.$$

This quantity is strictly equal to 1 for positive definite AFMC integrands (closed-shell system of particles interacting via an attractive force). For the Coulomb-repelling electrons the integrand is not positive definite, and the sign is, in general, a positive number smaller than 1. Positive definiteness of the integrand is important when Metropolis importance sampling is performed. Although we do not use a Metropolis algorithm (we sample the auxiliary fields directly from a Gaussian distribution), the sign Σ_0 is a useful indication of the stability of the Monte Carlo integrand. We generalize the sign concept to the shifted-contour case by

$$\Sigma_\alpha(\beta) = \frac{\left\langle e^{i\int_0^\beta \alpha^T(\tau) V \sigma(\tau) d\tau} U_{\sigma-i\alpha}^{(1)}(\beta) \right\rangle_W}{\langle |U_{\sigma-i\alpha}^{(1)}(\beta)| \rangle_W}. \tag{5.3}$$

The sign Σ_α is plotted as a function of the inverse temperature β in Fig. 5. When no contour shift is employed (the usual AFMC), the sign Σ_0 decreases exponentially fast to zero. However, when a contour shift using the Hartree–Fock density is made, the sign decreases only weakly with increasing β. The use of a closed-shell Hartree–Fock density at large bond length results in a degraded stabilization, as already pointed out in the discussion of Fig. 3.

VI. CONCLUSIONS AND DISCUSSION

We have shown that a considerable stabilization of the Monte Carlo integration can be achieved when the exact electron density is used to shift the Hubbard–Stratonovich line-integral contour into the complex plane. Unfortunately,
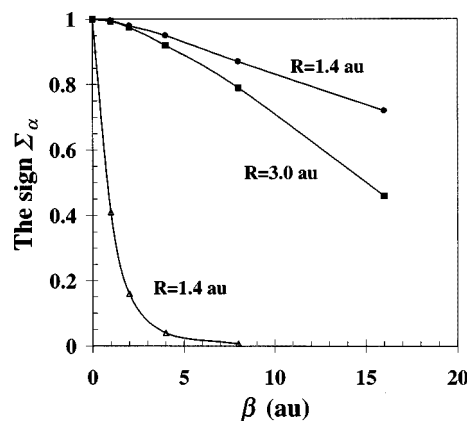
calculating the exact density is as computationally demanding as the full solution of the ground state electronic structure problem. The Hartree–Fock density can be used as an approximate stabilizer.[32] When this is done, a remarkable improvement is indeed seen in the statistical convergence of the ground state energy of H_2 and in the behavior of the sign Σ. In fact, within the range of β values we have tested (up to $\beta = 16$ a.u.), we have not seen evidence of an exponential increase of the statistical fluctuations or an exponential decrease of the sign.

The conclusion is that an overwhelming part of the ''sign problem'' is not due to electron–electron correlations and can be efficiently dealt with by using a mean-field density within the MC integration. This is somewhat analogous to the fact that the fixed nodes of the HF single determinant wave function within DMC can account for over 90% of the correlation energy of first row atoms and dimer.[16,45] It is also in accord with the good results obtained by using a HF determinant in VMC, and with the CPMC[15] which use an unrestricted Hartree-Fock trial function for imposing a determinant-manifold fixed node approximation. However, there is one important distinction that should be stressed: the AFMC integration is always a *formally exact procedure*. Contrarily, a fixed node procedure in DMC, GFMC, and CPMC is an approximation. The methods for systematically improving the fixed node approximations in DMC[46–49] are usually either severely limited by the sign problem or quickly become computationally prohibitive with system size.[17,19]

Stabilizing of AFMC is rather easily implemented in an already running AFMC code. It does not burden the calculation with significant additional numerical work (Hartree–Fock is anyway executed as a first step for getting the initial determinant). However, while showing great potential it is not yet clear at this point what are the practical limitations of this new AFMC method. Clearly more work is required.

ACKNOWLEDGMENTS

[1] W. L. McMillan, Phys. Rev. A **138**, 442 (1982).

[2] D. M. Ceperley, Phys. Rev. B **18**, 3126 (1978).

[3] S. B. Fahy, X. W. Wang, and S. G. Louie, Phys. Rev. Lett. **61**, 1631 (1980).

[4] M. H. Kalos, Phys. Rev. **128**, 560 (1962).

[5] M. H. Kalos, D. Levesque, and L. Verlet, Phys. Rev. A **9**, 2178 (1974).

[6] D. M. Ceperley and B. J. Alder, Phys. Rev. Lett. **45**, 566 (1980).

[7] D. M. Ceperley and M. H. Kalos, in *Monte Carlo Methods in Statistical Physics,* edited by K. Binder, 2nd ed. (Springer-Verlag, Berlin, 1986), p. 145.

[8] J. B. Anderson, J. Chem. Phys. **63**, 1499 (1975).

[9] P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester, Jr., J. Chem. Phys. **77**, 5593 (1982).

[10] D. M. Ceperley, Phys. Rev. Lett. **69**, 331 (1992).

[11] D. M. Ceperley, Rev. Mod. Phys. **67**, 279 (1995).

[12] G. Sugiyama and S. E. Koonin, Ann. Phys. (N.Y.) **168**, 1 (1986).

[13] S. B. Fahy and D. R. Hamann, Phys. Rev. Lett. **65**, 3437 (1990).

[14] S. Zhang, J. Carlson, and J. E. Gubernatis, Phys. Rev. Lett. **74**, 3652 (1995).

[15] S. Zhang, J. Carlson, and J. E. Gubernatis, Phys. Rev. B **55**, 7464 (1997).

[16] K. E. Schmidt and M. H. Kalos, in *Applications of the Monte Carlo Method in Statistical Physics,* edited by K. Binder, 2nd ed. (Springer-Verlag, Berlin, 1987), p. 125.

[17] B. L. Hammond, W. A. Lester, Jr., and P. J. Reynolds, *Monte Carlo Methods in Ab-Initio Quantum Chemistry* (World Scientific, Singapore, 1994).

[18] J. B. Anderson, Int. Rev. Phys. Chem. **14**, 85 (1995).

[19] D. M. Ceperley and L. Mitas, Adv. Chem. Phys. **93**, 1 (1996).

[20] D. M. Ceperley and B. Bernu, J. Chem. Phys. **89**, 6316 (1988).

[21] R. D. Stratonovich, Dokl. Akad. Nauk SSSR **115**, 1907 (1957); [translation Sov. Phys. Dokl. **2**, 416 (1958)].

[22] J. Hubbard, Phys. Rev. Lett. **3**, 77 (1959).

[23] S. Sorella, S. Baroni, R. Car, and M. Parrinello, Europhys. Lett. **8**, 663 (1989).

[24] S. Sorella, E. Tosatii, S. Baroni, R. Car, and M. Parinello, Int. J. Mod. Phys. B **1**, 993 (1988).

[25] J. E. Hirsch, Phys. Rev. B **31**, 4403 (1985).

[26] S. R. White, D. J. Scalapino, R. L. Sugar, E. Y. Loh, J. E. Gubernatis, and R. T. Scarlettar, Phys. Rev. B **40**, 506 (1989).

[27] D. R. Hamann and S. B. Fahy, Phys. Rev. B **41**, 11352 (1990).

[28] E. Y. Loh, Jr., J. E. Gubernatis, R. T. Scarlettar, S. R. White, D. J. Scalapeno, and R. L. Sugar, Phys. Rev. B **41**, 9301 (1990).

[29] D. M. Charutz and D. Neuhauser, J. Chem. Phys. **102**, 4495 (1995).

[30] P. L. Silvestrelli, S. Baroni, and R. Car, Phys. Rev. Lett. **71**, 1148 (1993).

[31] S. Fahy and D. R. Hamann, Phys. Rev. B **43**, 765 (1991).

[32] N. Rom, D. M. Charutz, and D. Neuhauser, Chem. Phys. Lett. **270**, 382 (1997).

[33] H. Kleinert, Phys. Lett. B **69**, 9 (1977).

[34] S. Levit, Phys. Rev. C **21**, 1594 (1980).

[35] M. C. Payne, M. P. Teter, D. C. Allen, T. A. Arias, and J. D. Joanopolus, Rev. Mod. Phys. **64**, 1045 (1992).

[36] R. Kosloff, in *Dynamics of Molecules and Chemical Reactions,* edited by R. E. Wyatt and J. Z. Zhang (Marcel Dekker, New York, 1996), p. 185.

[37] B. G. Johnson, P. M. W. Gill, M. Head-Gordon, C. A. White, J. Baker, D. R. Maurice, T. R. Adams, J. Kong, M. Challacombe, E. Schwegler, M. Oumi, C. Ochsenfeld, N. Ishikawa, J. Florian, R. D. Adamson, J. P. Dombroski, R. L. Graham, and A. Warshel, Q-Chem, Version 1.1, Q-Chem, Inc., Pittsburgh, PA, 1997.

[38] T. H. Dunning, Jr., J. Chem. Phys. **90**, 1007 (1989).

[39] A. J. Williamson, G. Rajagopal, R. J. Needs, L. M. Fraser, W. M. C. Foulkes, Y. Wang, and M.-Y. Chou, Phys. Rev. B **55**, 4851 (1997).

[40] L. M. Fraser, W. M. C. Foulkes, G. Rajagopal, R. J. Needs, S. D. Kenny, and A. J. Williamson, Phys. Rev. B **53**, 1814 (1996).

[41] G. Rajagopal, R. J. Needs, S. Kenny, W. M. C. Foulkes, and A. James, Phys. Rev. Lett. **73**, 1959 (1994).

[42] R. Kosloff, J. Phys. Chem. **92**, 2087 (1988).

[43] R. Kosloff, Annu. Rev. Phys. Chem. **45**, 145 (1994).

[44] D. K. Hoffman, Y. Huang, W. Zhu, and D. J. Kouri, J. Chem. Phys. **101**, 1242 (1994).

[45] R. P. Subramaniam, M. A. Lee, K. E. Schmidt, and J. W. Moskowitz, J. Chem. Phys. **97**, 2600 (1992).

[46] D. Ceperley and B. J. Adler, J. Chem. Phys. **81**, 5833 (1984).

[47] R. Bianchi, D. Bressanini, P. Cremaschi, and G. Morosi, J. Chem. Phys. **98**, 7204 (1993).

[48] J. B. Anderson, C. A. Traynor, and B. M. Boghosian, J. Chem. Phys. **95**, 7418 (1991).

[49] S. Zhang and M. H. Kalos, Phys. Rev. Lett. **67**, 3074 (1991).