Energy measurement in auxiliary-field many-electron calculations

D. R. Hamann and S. B. Fahy
AT&T Bell Laboratories, 600 Mountain Avenue, Murray Hill, New Jersey 07974
(Received 11 January 1990)

The ground state of a many-electron system can be projected from a trial state using an integral of one-electron projectors over a set of Hubbard-Stratonovich auxiliary fields. We have explored several alternative formulations of this method and find that the convergence of the ground-state energy with the number of field variables can be dramatically improved by choosing the appropriate one. This is illustrated using the two-dimensional Hubbard model as an example. These results are also used to examine the error associated with neglecting the so-called "sign problem" common to quantum Monte Carlo many-fermion calculations.

I. INTRODUCTION

Nonperturbative Monte Carlo methods offer the promise of systematically improvable calculations of electron correlation energies in atoms, molecules, and solids at a computational cost that grows at a reasonable rate with the complexity of the system. The most widely explored of these are variational methods in which a small number of parameters are varied to minimize an energy estimated by Monte Carlo evaluation of the expectation value of the many-electron Hamiltonian as a coordinate-space integral.[1-3] While that variational character of this approach is a distinct advantage, the small number of parameters that can practically be varied limits it's capabilities in situations in which we cannot guess a good wave function. To overcome this limitation, the "diffusion" and "Green's-function" Monte Carlo methods have been developed, in which an initial trial state is evolved towards the ground state by successive applications of the many-body Hamiltonian or its inverse.[2,4] This approach is also applied in coordinate space, and is limited by the "sign problem" produced by the antisymmetry of the electron wave function. The evolution procedures can only represent the exact antisymmetric ground state as a finite difference between diverging quantities, and the ground-state energy can only be estimated transiently.[4,5]

An alternative method of evolving a trial state towards the ground state was recently introduced by Sugiyama and Koonin,[6] and pursued by Sorella *et al.*[7,8] In this approach, the operator,

$$\hat{U}(\tfrac{1}{2}\beta)=e^{-\frac{1}{2}\beta\hat{H}}, \tag{1}$$

where β is regarded as an imaginary time or inverse temperature, serves to project out the ground state of $\hat{H}$ as $\beta\rightarrow\infty$. The transformation introduced by Stratonovich[9] and Hubbard[10] for the calculation of many-body partition functions is applied to $\hat{U}$, and the two-particle interaction term in $\hat{H}$ is replaced by one-particle interactions with a set of random time-varying auxiliary fields $\{x_i(\tau)\}$. Integration of $\hat{U}$ over a Gaussian distribution of these fields restores the interaction. The potential advantage of this approach is that for any particular member of this distribution, $\hat{U}[x_i(\tau)]$ is a single-particle operator, and a trial state chosen to be a Slater determinant remains a Slater determinant under propagation by this $\hat{U}$. The ground state is then built up as a sum of determinants, so antisymmetry is preserved term by term.

The auxiliary-field formulation of the finite temperature partition function has been used extensively to study electron correlations in the Hubbard model.[11] The development of the methods for carrying out these calculations has been dominated by a tendency to exploit the simplicity of the Hubbard model to maximize computational efficiency.[11,12] While the ground-state projection and grand canonical partition function approaches are closely related, there are enough differences in both computational detail and physical interpretation to motivate a fresh look at the methodology. We have used the Hubbard model as a convenient test case in such a pursuit. However, our long-range goal is to develop these methods for real materials, so we have generally ignored the model-specific simplifications. In so doing, we have discovered that nominally equivalent approaches to measuring the energy have substantially different rates of convergence with the number of auxiliary-field variables. We can reproduce the rate reported for the well-developed approaches.[11] However, we find that two alternative formulations give 1 and 2 orders of magnitude better convergence, respectively. We can at least rationalize this improvement in terms of a picture which we are developing of the behavior of the auxiliary fields and the response of the electrons.

This study also addresses a second issue. Despite the fact that the auxiliary-field method preserves explicitly the antisymmetric structure of the electron wave function under particle interchange, the energy formally emerges as a difference between diverging quantities as $\beta\rightarrow\infty$ (except in special cases). Arguments and numerical evidence have been presented indicating that the average rather than the difference of these quantities gives an accurate estimator of the energy.[7,8] These arguments have been disputed on the basis of other numerical evidence.[13] We have chosen as examples two cases for which the ground-state energy is known from exact calculations, and found that the error of the "average" approach lies

well outside our statistical uncertainty. Our study of this error motivates the introduction of yet another approach to energy measurement. We will report an extended study of other aspects of the "sign problem" separately.[14]

The remainder of this paper is organized as follows: Section II describes the methodology in a self-contained manner, distinguishing the four approaches to energy measurement and discussing our approach to sampling the auxiliary fields. Section III presents and discusses our ground-state energy results for the 2×2 two-dimensional Hubbard model with three electrons, and the 4×4 with 14 electrons. Following our conclusions in Sec. IV, we present simple derivations of several key formulas, and the relation between the projection and grand canonical methods in the Appendix.

II. METHODS

A. Auxiliary-field projection and measurement

As discussed in the introduction, we wish to use the operator $\hat{U}(\beta)$ defined in Eq. (1) to project the many-electron ground state from a trial state. (In general, we will use capital letters with hats to denote operators on many-electron wave functions.) The trial state, in a notation that will be useful for further development, is

$$|\Phi_0\rangle=[\mathbf{c}^\dagger*\boldsymbol{\phi}_1(0)][\mathbf{c}^\dagger*\boldsymbol{\phi}_2(0)]\cdots[\mathbf{c}^\dagger*\boldsymbol{\phi}_M(0)]|0\rangle\ , \tag{2}$$

where $\mathbf{c}^\dagger$ is a row vector of creation operators $c_i^\dagger$ in an orthogonal one-electron basis (a site basis in the case of the Hubbard model), $\boldsymbol{\phi}_j$ is a column vector of the coefficients ϕ_{ij} of the jth one-electron wave function in the M-electron trial state, and $|0\rangle$ is the electron vacuum. A scalar product of $\mathbf{c}^\dagger$ and $\boldsymbol{\phi}_j$ over the one-electron basis is implied by an asterisk, a convention which we will use throughout. We will generally suppress the spin index, which could be considered as a component of the basis index i, and will explicitly write the pair (i,σ) when we need to express it. The projected state is

$$|\Phi_{\beta/2}\rangle=\hat{U}(\tfrac{1}{2}\beta)|\Phi_0\rangle\ , \tag{3}$$

and the ground-state energy is the large β limit of

$$E(\beta)=\frac{\langle\Phi_{\beta/2}|\hat{H}|\Phi_{\beta/2}\rangle}{\langle\Phi_{\beta/2}|\Phi_{\beta/2}\rangle}=\frac{\langle\Phi_0|e^{-\frac{1}{2}\beta\hat{H}}\hat{H}e^{-\frac{1}{2}\beta\hat{H}}|\Phi_0\rangle}{\langle\Phi_0|e^{-\beta\hat{H}}|\Phi_0\rangle}\ . \tag{4}$$

Since $\hat{H}$ commutes with any function of itself, we can split either projection operator into two pieces and place $\hat{H}$ in between. An equivalent expression for $E(\beta)$ is obtained by averaging over all possible ways of so doing,

$$E(\beta)=\frac{\int_0^\beta d\tau\langle\Phi_0|e^{-(\beta-\tau)\hat{H}}\hat{H}e^{-\tau\hat{H}}|\Phi_0\rangle}{\beta\langle\Phi_0|e^{-\beta\hat{H}}|\Phi_0\rangle}\ . \tag{5}$$

This latter form gives better statistics for measuring $\hat{H}$ when we carry out the auxiliary-field integration by stochastic methods. This form also reveals our motivation for introducing $\beta/2$ rather than simply β in the previous definitions.

The Hubbard model we shall use as an example has the Hamiltonian

$$\hat{H}=-\sum_{i,j=1}^{N}T_{ij}c_i^\dagger c_j+U_{\rm Hub}\sum_{i=1}^{N}n_{i\uparrow}n_{i\downarrow}=-\mathbf{c}^\dagger*\underline{T}*\mathbf{c}+U_{\rm Hub}\mathbf{n}_\uparrow^\dagger*\mathbf{n}_\downarrow\ , \tag{6}$$

where we have extended our vector notation to the electron number operators. We will take the kinetic-energy matrix to represent nearest-neighbor hopping with unit magnitude for a periodically extended two-dimensional lattice, but nothing in our treatment exploits this sparseness. Following recent practice, we will rearrange the interaction term to contain the square of the site magnetization,

$$U_{\rm Hub}\mathbf{n}_\uparrow^\dagger*\mathbf{n}_\downarrow=-\tfrac{1}{2}U_{\rm Hub}(\mathbf{n}_\uparrow-\mathbf{n}_\downarrow)^\dagger*(\mathbf{n}_\uparrow-\mathbf{n}_\downarrow)+\tfrac{1}{2}U_{\rm Hub}\hat{N}\ , \tag{7}$$

where $\hat{N}$ is the total electron number operator. Other rearrangements, such as one symmetric in charge and magnetization, have proven useful,[15] but will not be considered here. Introducing the Feynman time-ordering operator T acting on the imaginary time τ, we can carry out the Hubbard-Stratonovich transformation[9,10] and write the projection operator as

$$\hat{U}(\beta)=\int\delta\mathbf{x}(\tau)T\exp\left[-\int_0^\beta d\tau\{\tfrac{1}{2}\mathbf{x}^t(\tau)*\mathbf{x}(\tau)-\mathbf{c}^\dagger(\tau)*T*\mathbf{c}(\tau)+\sqrt{U_{\rm Hub}}\mathbf{x}^t*[\mathbf{n}_\uparrow(\tau)-\mathbf{n}_\downarrow(\tau)]\}\right]\ . \tag{8}$$

In the above, $\mathbf{x}(\tau)$ is the set of Hubbard-Stratonovich auxiliary fields, which is a vector in the site indices and a function of the continuous variable τ. The measure of the functional integral $\delta\mathbf{x}$ is assumed to contain the appropriate normalization, which can be defined as a limit but is irrelevant here. The $-\frac{1}{2}U_{\rm Hub}\hat{N}$ term in Eq. (7) represents a constant offset of the single-particle energy scale, and is subsumed in the kinetic energy matrix T. While the $\mathbf{x}(\tau)$ are explicitly τ-dependent functions, the electron operators $\mathbf{c}^\dagger$, $\mathbf{c}$, and $\mathbf{n}_\sigma$ have τ arguments simply to indicate their ordering under the action of T.

Defined in this manner, the Hubbard-Stratonovich transformation is exact. It has become customary to introduce a finite time slice $\Delta\tau$ and use the Trotter approximation[16] to separate the noncommuting kinetic and interaction terms prior to carrying out the transformation.[7,8] The Feynman ordering approach used in the original derivation by Hubbard[10] is perfectly valid, how-

ever, and of course the two are equivalent in the limit $\Delta\tau \to 0$.

We have belabored the distinction between Eq. (8) and the Trotter form because it has important consequences for our results. Having made this point, we now introduce a stepwise approximation for $\mathbf{x}(\tau)$,

$$\mathbf{x}(\tau)=\mathbf{x}_l, \quad (l-1)\Delta\tau<\tau<l\Delta\tau, \quad l=1,P \ , \tag{9}$$

where $\beta=P\Delta\tau$, and $\{\mathbf{x}_l\}$ is a discrete set of auxiliary fields. Substituting Eq. (9) in Eq. (8),

$$\hat{U}=\int_{-\infty}^{\infty} d\mathbf{x}_1 \cdots d\mathbf{x}_P \exp\left[\frac{-\Delta_\tau}{2}\sum_{l=1}^{P}\mathbf{x}_l^{\dagger}*\mathbf{x}_l\right]\hat{U}_P\cdots\hat{U}_2\hat{U}_1$$
$$=\int \delta\mathbf{x} G\hat{U}_P\cdots\hat{U}_2\hat{U}_1 \ , \tag{10}$$

where we have used G to denote the Gaussian weighting factor and $\hat{U}$ with a subscript l to denote the time-slice propagators,

$$\hat{U}_l=\exp(-\Delta\tau\hat{H}_l) \ , \tag{11}$$

where

$$\hat{H}_l=\hat{T}+\hat{V}_l$$
$$=-\mathbf{c}^{\dagger}*T*\mathbf{c}+\sqrt{U_{\mathrm{Hub}}}\mathbf{x}_l^{\dagger}*(\mathbf{n}_{\downarrow}-\mathbf{n}_{\uparrow}) \ . \tag{12}$$

The time-slice Hamiltonian $\hat{H}_l$ is τ independent over the time slice. While we have restricted our functional integration over $\mathbf{x}(\tau)$ through Eq. (9), we have not made any assumptions about the commutativity of $\hat{T}$ and $\hat{V}_l$. If we wish, we can now make the Troter approximation,[16]

$$\hat{U}_l=\exp(-\tfrac{1}{2}\Delta\tau\hat{T})\exp(-\Delta\tau\hat{V}_l)\exp(-\tfrac{1}{2}\Delta\tau\hat{T})+O(\Delta\tau^3) \ . \tag{13}$$

If we had introduced $\Delta\tau$ and Trotter prior to the Hubbard-Stratonovich transformation, we would of course have obtained the form in Eq. (13) directly. While we had no *a priori* argument to show that Eq. (11) should be any better than Eq. (13), we have discovered numerically that it is, and can rationalize our observation.

The time-slice propagators operate on the trial state $|\Phi_0\rangle$. Using Eq. (A5) we find

$$\hat{U}_1|\Phi_0\rangle=e^{-\Delta\tau\hat{H}_1}[\mathbf{c}^{\dagger}*\boldsymbol{\phi}_1(0)][\mathbf{c}^{\dagger}*\boldsymbol{\phi}_2(0)]\cdots[\mathbf{c}^{\dagger}*\boldsymbol{\phi}_M(0)]|0\rangle$$
$$=[\mathbf{c}^{\dagger}*e^{-\Delta\tau\underline{H}_1}*\boldsymbol{\phi}_1(0)]e^{-\Delta\tau\hat{H}_1}[\mathbf{c}^{\dagger}*\boldsymbol{\phi}_2(0)]\cdots[\mathbf{c}^{\dagger}*\boldsymbol{\phi}_M(0)]|0\rangle$$
$$=[\mathbf{c}^{\dagger}*e^{-\Delta\tau\underline{H}_1}*\boldsymbol{\phi}_1(0)][\mathbf{c}^{\dagger}*e^{-\Delta\tau\underline{H}_1}*\boldsymbol{\phi}_2(0)]\cdots[\mathbf{c}^{\dagger}*e^{-\Delta\tau\underline{H}_1}*\boldsymbol{\phi}_M(0)]|0\rangle \ , \tag{14}$$

where $e^{-\Delta\tau\underline{H}_l}\equiv\underline{U}_l$ is a matrix in the one-electron basis, as opposed to the Fock-space operator $\hat{U}_l$. We can apply each operator in $\hat{U}_j$ in turn for $j=1,2,\ldots,l$ and the same steps show that we simply accumulate a product of the matrices $\underline{U}_j$ operating on each single-particle wave function $\boldsymbol{\phi}_k$. The result is the "advanced" Slater determinant

$$|\Phi_l^{>}\rangle=[\mathbf{c}^{\dagger}*\boldsymbol{\phi}_1^{>}(l)][\mathbf{c}^{\dagger}*\boldsymbol{\phi}_2^{>}(l)]\cdots[\mathbf{c}^{\dagger}*\boldsymbol{\phi}_M^{>}(l)]|0\rangle \ , \tag{15}$$

where

$$\boldsymbol{\phi}_k^{>}(l)=\underline{U}_l*\underline{U}_{l-1}*\cdots*\underline{U}_2*\underline{U}_1*\boldsymbol{\phi}_k(0) \ . \tag{16}$$

As we see from Eqs. (15) and (16), each one-electron state of the trial Slater determinant is propagated independently through the τ-varying field $\sqrt{U_{\mathrm{Hub}}}\mathbf{x}(\tau)$ given by Eq. (9). In practice, when each $\underline{U}_j$ operates on the set $\{\boldsymbol{\phi}_k(j-1)\}$, all the wave functions in the resulting new set tend to become parallel to the eigenvector of $\underline{U}_j$ with the largest eigenvalue. Thus the straightforward application of Eq. (16) will result in a Slater determinant with a great deal of linear dependence among the wave functions. This dependence can easily exhaust the precision of floating-point arithmetic in the computational application of these methods. Sugiyama and Koonin were the first to recognize this problem. They proposed the practical solution of propagating the entire set of $\boldsymbol{\phi}$'s together, and applying Schmidt orthogonalization after every few time slices.[6] We orthonormalize, typically after each time slice, and accumulate the log of the norm, since the norm itself can outrun the dynamic range of double-precision arithmetic for large β. We will not introduce any extra notation to denote the effects of the orthogonalization steps, since they play no real role in the subsequent development.

We will also need to allow the time-slice propagators to operate to the left, so we introduce the "retarded" wave functions

$$\boldsymbol{\phi}_k^{<}(l)=\underline{U}_{l+1}^{\dagger}*\cdots*\underline{U}_{P-1}^{\dagger}*\underline{U}_P^{\dagger}*\boldsymbol{\phi}_k(0) \ . \tag{17}$$

They are defined in such a way that

$$\langle\Phi_0|\hat{U}_P\cdots\hat{U}_2\hat{U}_1|\Phi_0\rangle=\langle\Phi_l^{<}|\Phi_l^{>}\rangle \tag{18}$$

for any l. If we define $M\times M$ overlap matrix of the retarded and advanced wave functions at l as

$$S_{ij}^l=\boldsymbol{\phi}_i^{<\dagger}(l)*\boldsymbol{\phi}_j^{>}(l) \ , \tag{19}$$

then, using (A6),

$$\langle\Phi_l^{<}|\Phi_l^{>}\rangle=\det_{ij}[S_{ij}^l]=D(\{\mathbf{x}_k\}) \ . \tag{20}$$

The final notation stresses that D is a function of the auxiliary fields but not of l. It can be seen from Eqs. (16), (17), and (19) that S_{ij}^l is independent of l in any range of l in which we do not insert Schmidt orthogonalization steps. It could be extended beyond such regions by triangular matrix transformations corresponding to the Schmidt orthogonalizations,[11] but this is of no practical

use in the context of projection method calculations.

To complete the discussion of energy measurement strategies, we must consider the Trotter breakup of the time-slice propagator as given in Eq. (13). Figuratively speaking, the time slice is divided into thirds, and we introduce the suggestively labeled wave functions corresponding to propagation through these sections,

$$\begin{aligned}
\phi_j^{>}(l+\tfrac{1}{3}) &= e^{-\frac{1}{2}\Delta\tau T} * \phi_j^{>}(l) , \\
\phi_j^{>}(l+\tfrac{2}{3}) &= e^{-\Delta\tau V_{l+1}} * \phi_j^{>}(l+\tfrac{1}{3}) , \\
\phi_j^{<}(l+\tfrac{2}{3}) &= (e^{-\frac{1}{2}\Delta\tau T})^t * \phi_j^{<}(l+1) , \\
\phi_j^{<}(l+\tfrac{1}{3}) &= (e^{-\Delta\tau V_{l+1}})^t * \phi_j^{<}(l+\tfrac{2}{3}) .
\end{aligned} \tag{21}$$

When the Trotter breakup is used, these definitions are such that Eqs. (18) and (20) continue to be obeyed for these fractional l values.

Three of the approaches we used to calculate the energy are based on Eq. (5), discretized to finite $\Delta\tau$. Using this, Eq. (10), and the preceding wave function definitions, we have

$$E(\beta) = \frac{\int \delta\mathbf{x}\, G \sum_{l=0}^{P} w_l \langle \Phi_l^{<} | \hat{H} | \Phi_l^{>} \rangle}{P \int \delta\mathbf{x}\, G D(\mathbf{x})} , \tag{22}$$

where $\hat{H}$ is the full, untransformed Hubbard Hamiltonian given in Eq. (6) and w_l is $\frac{1}{2}$ for $l=0$ or $l=P$, and 1 otherwise. While the w_l factors resemble the trapezoidal integration rule, they actually represent the symmetric disposition of the energy measurement on each side of the time-slice operator, $\frac{1}{2}(\hat{H}\hat{U}_l + \hat{U}_l\hat{H})$. This is nontrivial, since these operators do not commute inside the $\delta\mathbf{x}$ integration. The temptation to replace the "trapezoidal rule" weights by a "better" integration formula must be avoided, since $\mathbf{x}(\tau)$ is not a continuous function.

We have not dwelt on spin in our development so far. While it could be lumped with the site index, it is more efficient to recognize that the $\hat{H}_l$ do not mix spin up and spin down, so our wave functions can be regarded as products of two Slater determinants, one for each spin. Using Eq. (A9), the matrix elements of $\hat{H}$ can be expressed as

$$\langle \Phi_l^{<} | \hat{H} | \Phi_l^{>} \rangle = \mathrm{Det}_{ij}[S_{ij}^{\uparrow l}]\mathrm{Det}_{ij}[S_{ij}^{\downarrow l}] \left[\langle T \rangle^{\uparrow l} + \langle T \rangle^{\downarrow l} + U_{\mathrm{Hub}} \sum_{k=1}^{N} \langle n_k \rangle^{\uparrow l} \langle n_k \rangle^{\downarrow l} \right] , \tag{23}$$

where

$$\langle T \rangle^{\sigma,l} = \sum_{ij=1}^{M} [(S^{\sigma,l})^{-1}]_{ij} \phi_j^{\sigma <}(l) * \underline{T} * \phi_i^{\sigma >}(l) \tag{24}$$

and

$$\langle n_k \rangle^{\sigma,l} = \sum_{ij=1}^{M} [(S^{\sigma,l})^{-1}]_{ij} \phi_{kj}^{\sigma <}(l) \phi_{ki}^{\sigma >}(l) . \tag{25}$$

Note that we have introduced components of the wave functions, ϕ_{ki} with site index k in Eq. (25), since the U_{Hub} term cannot be clearly expressed with the shorthand vector notation.

Equations (22)–(24) can be applied in two ways. If we evaluate Eq. (11) for the time-slice propagator $\hat{U}_l$ exactly, we will denote the corresponding energy by $E_1(\beta)$. If we use the Trotter approximation as given in Eq. (13) for $\hat{U}_l$, we will denote the corresponding energy by $E_2(\beta)$. Of course they must become identical as $\Delta\tau \to 0$.

We will discuss efficient means to sample the $\delta\mathbf{x}$ integral in the next section. We note here that when using the Trotter breakup, our procedure will require evaluating the derivatives of $\exp(-\Delta\tau\hat{V}_l)$ with respect to x_{kl}. These are proportional to $\langle n_k \rangle^{\sigma,l+1/3}$, using the "fractional time-slice" notation introduced in Eq. (21), (or equivalently to $\langle n_k \rangle^{\sigma,l+2/3}$, since the $\hat{n}_k$ commute). It is then computationally economical to measure the energy at the same points. Since $\hat{T}$ does not commute with $\hat{V}_l$, the symmetric formula

$$E_3(\beta) = \left[P \int \delta\mathbf{x}\, G D(\mathbf{x}) \right]^{-1} \left[\int \delta\mathbf{x}\, G \sum_{l=0}^{P-1} \tfrac{1}{2} (\langle \Phi_{l+1/3}^{<} | \hat{H} | \Phi_{l+1/3}^{>} \rangle + \langle \Phi_{l+2/3}^{<} | \hat{H} | \Phi_{l+2/3}^{>} \rangle) \right] \tag{26}$$

seems to be the best choice.[17] While this appears to double the calculational effort for the kinetic-energy calculation, it does not. It can easily be shown that

$$\langle T \rangle^{\sigma,l+1/3} + \langle T \rangle^{\sigma,l+2/3} = \sum_{ij=1}^{M} \left[[(S^{\sigma,l+2/3})^{-1}]_{ij} \times \sum_{kn=1}^{N} \phi_{kj}^{\sigma <}(l+\tfrac{2}{3}) T_{kn} \phi_{ni}^{\sigma >}(l+\tfrac{2}{3}) (1 + e^{-\Delta\tau\sigma\sqrt{U_{\mathrm{Hub}}}(x_{kl} - x_{nl})}) \right] . \tag{27}$$

The sums here and in Eq. (24) can be viewed as a density matrix traced with T_{kn}. This matrix can be accumulated over the l index and traced over the k,n indices only once. In the $\Delta\tau\to 0$ limit, $E_3(\beta)$ must converge to the same value as $E_1(\beta)$ and $E_2(\beta)$, and we have no *a priori* arguments establishing which should converge more rapidly.

B. Sampling the auxiliary fields

The $\delta\mathbf{x}$ integral to be performed in evaluating either Eq. (22) or Eq. (26) for the energy has far too high a dimension to evaluate by any mesh approach, and some importance sampling scheme must be used. These integrals all have the schematic form

$$E(\beta)=\frac{\int \delta\mathbf{x}\, GDE}{\int \delta\mathbf{x}\, GD}\ , \tag{28}$$

where E in the integrand (as well as G and D, of course) is understood to be a functional of $\mathbf{x}$. It is tempting to treat GD as a probability, but D is positive only in special cases. The standard approach is to use $G|D|$ as the probability to be sampled,[7,8,11] so

$$E(\beta)=\frac{\int \delta\mathbf{x}\, G|D|sE}{\int \delta\mathbf{x}\, G|D|s}=\frac{\int \delta\mathbf{x}\, G|D|sE}{\langle s\rangle}\ , \tag{29}$$

where the $s=D/|D|$ is the sign, and $\langle s\rangle$ is its average over the probability distribution.

Standard Monte Carlo sampling has been used most extensively in the application of the auxiliary-field method to the Hubbard model. However, this is usually done within the context of the discrete Hubbard-Stratonovich transformation, which is specialized to this model.[12] We have retained the continuous version because of our interest in generalizing to realistic systems, and it is difficult to sample the continuous $\mathbf{x}$ fields efficiently by the standard Monte Carlo approach. We have therefore adopted the hybrid Monte Carlo scheme of Duane *et al.*, which was introduced to deal with similar problems in simulations of lattice field theory.[18] Briefly, a set of momenta considered to be conjugate to the $\mathbf{x}$'s is introduced, and the numerator and denominator of Eq. (29) are multiplied by a Gaussian integral over these momenta (velocities, actually, since we take unit mass),

$$E(\beta)=\frac{\int \delta\mathbf{v}\, G_v \int \delta\mathbf{x}\, G|D|sE}{\int \delta\mathbf{v}\, G_v \int \delta\mathbf{x}\, G|D|s}\ , \tag{30}$$

where

$$G_v=\exp\left[\frac{-\Delta\tau}{2}\sum_{l=1}^{P}\mathbf{v}_l^t*\mathbf{v}_l\right]\ . \tag{31}$$

Integrating the probability $G_vG|D|$ over $\delta\mathbf{v}$ and $\delta\mathbf{x}$ is equivalent to evaluating a classical partition function at unit temperature for a set of NP particles moving under the action of the potential $-\ln(G|D|)$. If we have some initial field configuration $\mathbf{x}_I$ and we pick a random set of velocities distributed according to their Gaussian distribution $\mathbf{v}_I$, then all points along the classical trajectory determined by these initial conditions have equal probability. This is true simply because the classical energy $E^{\rm cl}=-(\ln G_vG|D|)$ is conserved along the trajectory. If we integrate the equations for an adequately long time Δt, we will generate a final field configuration $\mathbf{x}_F$ which will be uncorrelated with $\mathbf{x}_I$. By using $\mathbf{x}_F$ as our new $\mathbf{x}_I$ and choosing a new set of velocities, we can continue the process, and summing E in the numerator integrand in Eq. (30) evaluated at the series of field configurations generated in this manner will give the correctly weighted average.

There are two things wrong with this procedure. First, we cannot numerically integrate the equations of motion exactly, and thus can't conserve $E^{\rm cl}$ exactly. Second, if we could integrate them exactly, $\mathbf{x}$ would be confined to one region of its NP-dimensional space by the zeros of D, since the potential $-\ln G|D|$ would have logarithmic infinities at these locations. This would invalidate our sample. These problems are both elegantly solved by the following two steps: First, we use an integration algorithm that preserves classical phase-space volume exactly even though it may make substantial other errors. Second, we incorporate a Metropolis rejection step based on

$$p=\exp[E^{\rm cl}(\mathbf{x}_I,\mathbf{v}_I)-E^{\rm cl}(\mathbf{x}_F,\mathbf{v}_F)]\ .$$

If $p\geq 1$ we accept $\mathbf{x}_F$ as a new sample point, and if $p<1$, we accept it with probability p. If we reject it, we reuse $\mathbf{x}_I$ as a sample point, choose a new set of random velocities $\mathbf{v}_I$, and integrate again.[18]

The preceding procedure generates a set of points which correctly samples the desired distribution.[18] We have considerable latitude in applying it since we can choose both the integration step length Δt and the number of molecular dynamics (MD) steps $n_{\rm MD}$ between Monte Carlo rejection steps. The effective potential for the classical problem tends to have the form of isolated favorable "valleys" separated by large unfavorable regions, so increasing error in the integration almost always makes $E_F^{\rm cl}-E_I^{\rm cl}$ more positive, and lowers our acceptance rate. The error increases when either Δt or $n_{\rm MD}$ is increased, while the statistical independence of $\mathbf{x}_I$ and $\mathbf{x}_F$ increases when the product $n_{\rm MD}\Delta t$ is increased. The Gaussian factors contribute independent harmonic-oscillator terms in the classical Hamiltonian which have a period 2π. This sets a natural time scale which is not greatly modified, on the average, by the contributions from the determinants. We find that taking $n_{\rm MD}\Delta t\sim 2.5$ gives good statistical independence ($\mathbf{x}_I*\mathbf{x}_F/|\mathbf{x}_I||\mathbf{x}_F|\sim 0.1$ on the average). We then pick Δt for an acceptance rate ~ 0.5, which is achieved with Δt in the range 0.2 to 0.4 for the cases we have investigated. The logarithmic barriers are very "thin" on the scale of the corresponding Δx_{kl}'s, and our procedure readily makes the "error" of going through them. The acceptance rate associated with these errors is no worse than that associated with the "valley walls" away from the sign-change infinities, and the rate of sign change with number of

Monte Carlo steps appears to be set by the time it takes the system to "find" these barriers.

The $\ln|D|$ term in the effective classical potential contributes forces which are long ranged in both lattice separation and τ. When we use the Trotter breakup of the $\hat{U}_l$'s, Eq. (13), it is easy to show that the force on x_{kl},

$$f_{kl}=\frac{\partial}{\partial x_{kl}}\ln\langle\Phi_0|\hat{U}_P\cdots\hat{U}_2\hat{U}_1|\Phi_0\rangle$$
$$=-\Delta\tau\sqrt{U_{\rm Hub}}(\langle n_k\rangle^{\uparrow,l+1/3}-\langle n_k\rangle^{\downarrow,l+1/3}) , \quad (32)$$

where the n_k expectation values are calculated according to Eq. (25).

When we use the exact operator for the time-slice propagator, as given by Eqs. (11) and (12), the force expression is not so obvious. To use the exact operator, we must first solve an eigenvalue problem for each time slice,

$$\underline{H}_l*\psi_{il}=(\underline{T}+\underline{V}_l)*\psi_{il}=\varepsilon_{il}\psi_{il} , \quad (33)$$

and then expand the one-electron propagation operator as

$$\underline{U}_l=\sum_{i=1}^{N}\psi_{il}\exp(-\Delta\tau\varepsilon_{il})\psi_{il}^{\dagger} . \quad (34)$$

Using Eq. (20), the force can be expressed as

$$\frac{\partial}{\partial x_{kl}}\ln\langle\Phi_l^{<}|\Phi_l^{>}\rangle=\frac{\partial}{\partial x_{kl}}\ln\mathrm{Det}_{ij}[S_{ij}^l]$$
$$=(\mathrm{Det}_{ij}[S_{ij}^l])^{-1}\sum_{ij}s_{ij}^l\frac{\partial}{\partial x_{kl}}S_{ij}^l$$
$$=\sum_{ij}[(S^l)^{-1}]_{ji}\frac{\partial}{\partial x_{kl}}S_{ij}^l , \quad (35)$$

where s_{ij}^l is the cofactor (signed minor) of S_{ij}^l in the expansion of the determinant, and we have used Cramers rule to replace the cofactor over determinant ratio with the inverse. To isolate the x_{kl} dependence of S^l, we rewrite Eq. (19) as

$$S_{ij}^l=\phi_i^{<\dagger}(l)*\underline{U}^l*\phi_j^{>}(l-1) . \quad (36)$$

Only $\underline{U}^l$ depends on x_{kl}. To find its derivative, we must recognize that derivatives of both ε_{il} and ψ_{il} in Eq. (34) are required. Both may be evaluated using first-order perturbation theory, and the resulting expression is

$$\frac{\partial}{\partial x_{kl}}\underline{U}_l=\sum_{ij=1}^{N}\psi_{jl}\left[\psi_{jl}^{\dagger}*\frac{\partial}{\partial x_{kl}}\underline{H}_l*\psi_{il}\right]$$
$$\times\frac{e^{-\Delta\tau\varepsilon_{il}}-e^{-\Delta\tau\varepsilon_{jl}}}{\varepsilon_{il}-\varepsilon_{jl}}\psi_{il}^{\dagger} . \quad (37)$$

When $\varepsilon_{il}\sim\varepsilon_{jl}$ or $i=j$ in the sum, the ratio should be replaced by its proper limiting value $-\Delta\tau e^{-\Delta\tau\varepsilon_{il}}$. Substituting $\sigma\sqrt{U_{\rm Hub}}n_{k\sigma}$ for $\partial\underline{H}_l/\partial x_{kl}$ in Eq. (37), the result into Eq. (36), and that expression into Eq. (35) gives the complete formula for the force when the exact propagator is used. The relationship of this formula to the physically transparent force expression for the Trotter approximation given in Eq. (32) is not obvious. However, one can show that Eqs. (35) and (37) are equivalent to calculating the average of $n_{k\sigma}$ inside the lth time slice.

While use of the exact $\hat{U}_l$ seems to produce considerable computational complexity, the numerical operations count in fact is proportional to N^3P. The Schmidt orthogonalization steps require of order NM^2P operations, and the propagation steps N^2MP operations (unless sparseness can be exploited). Since $M\sim N$ for Hubbard models, the limiting scale of the calculation is not changed. For realistic models, where the wave function basis size might be large compared to the number of electrons, the difference between N^2M and N^3 might be more significant.

C. The sign problem and derivative measurement

Since both the numerator and denominator in our energy expression, Eq. (29), are to be evaluated by a statistical sampling approach, it is clear that it will be difficult to obtain accurate statistics if $\langle s\rangle$ is small. This is the Fermion sign problem discussed in the Introduction. It has recently been argued that the quantity derived from Eq. (29) by setting $s=1$,

$$E^*(\beta)=\frac{\int\delta\mathbf{x}G|D|E}{\int\delta\mathbf{x}G|D|} , \quad (38)$$

converges to $E(\beta)$ as $\beta\rightarrow\infty$, as long as $\langle s(\beta)\rangle$ does not go to zero exponentially at large β.[7,8] Sorella *et al.* presented numerical examples showing that the $\langle s\rangle$ remained finite in cases where it was not forced to be unity by a special symmetry[8] (electron-hole symmetry in the half-filled Hubbard model on a bipartite lattice[14,19]). They also presented E^* results, which agreed within statistical accuracy with ground-state energies calculated by exact diagonalization. On the other hand, it has been shown by independent numerical examples that $\langle s(\beta)\rangle$ does go to zero as $\exp(-\Delta E\beta)$.[13] Loh *et al.* also present a heuristic argument that $E=E^*+\Delta E$, but do not illustrate this formula with any numerical examples.[13]

An energy correction of this form can be derived as follows. First, we rewrite our basic definition of $E(\beta)$, Eq. (4), as

$$E(\beta)=-\frac{d}{d\beta}\ln\langle\Phi_0|e^{-\beta\hat{H}}|\Phi_0\rangle$$
$$=-\frac{d}{d\beta}\ln\int\delta\mathbf{x}GD \quad (39)$$

Multiplying and dividing by the corresponding expression with D replaced by $|D|$ gives the desired result,

$$E(\beta)=-\frac{d}{d\beta}\ln\left[\int\delta\mathbf{x}\,G|D|\frac{\int\delta\mathbf{x}\,GD}{\int\delta\mathbf{x}\,G|D|}\right]$$
$$=E^*-\frac{d}{d\beta}\ln\langle s(\beta)\rangle . \quad (40)$$

There is a subtlety in the derivation of Eq. (40), however. In our other energy expressions, based on Eq. (5), we kept the full many-body $\hat{H}$ as the quantity we measured. This is equivalent to taking the β derivative in Eq. (39)

before carrying out the Hubbard-Stratonovich transformation. But E^* is defined only after the transformation, since positive and negative regions of $D(\mathbf{x})$ have no meaning in terms of many-body eigenstates.

To derive the strictly consistent expression for E^*, suppose we have carried out the transformation and introduced the discrete $\mathbf{x}(\tau)$ before taking the β derivative. Then $d/d\beta$ becomes $(1/P)d/d(\Delta\tau)$. To obtain an expression as close to our previous measures as possible, we change auxiliary-field variables to $y=\sqrt{\Delta\tau}x$ to remove the $\Delta\tau$ dependence from the Gaussian factor. We then find

$$E_4^*(\beta)=\frac{\int \delta\mathbf{y}\, Gs(\mathbf{y}) \sum_{l=1}^{P} \left\langle \Phi_l^< \left| \frac{\partial \hat{U}_l}{\partial(\Delta\tau)} \right| \Phi_{l-1}^> \right\rangle}{P\int \delta\mathbf{y}\, G|D(\mathbf{y})|}, \quad (41)$$

where we have introduced the subscript 4 to distinguish this from our other measures. The appearance of the sign $s(\mathbf{y})$ in the numerator is the result of carrying the derivative inside the absolute value. The matrix elements in the summand can be evaluated following the same approach as used to calculate the forces acting on the auxiliary fields. In analogy with Eq. (35), we find

$$\left\langle \Phi_l^< \left| \frac{\partial \hat{U}_l}{\partial(\Delta\tau)} \right| \Phi_{l-1}^> \right\rangle = D\sum_{ij} [(S^l)^{-1}]_{ji} \boldsymbol{\phi}_i^{<\dagger}(l) * \frac{\partial \underline{U}^l}{\partial(\Delta\tau)} * \boldsymbol{\phi}_j^>(l-1) . \quad (42)$$

The expression analogous to Eq. (37) for the derivative of the one-electron operator required in the preceding expression is

$$\frac{\partial \underline{U}^l}{\partial(\Delta\tau)} = \sum_{ij=1}^{N} \boldsymbol{\psi}_{jl}[\boldsymbol{\psi}_{jl}^\dagger * (\underline{T}+\tfrac{1}{2}\underline{V}_l) * \boldsymbol{\psi}_{il}] \times \frac{e^{-\Delta\tau\varepsilon_{il}} - e^{-\Delta\tau\varepsilon_{jl}}}{\Delta\tau(\varepsilon_{il}-\varepsilon_{jl})} \boldsymbol{\psi}_{il}^\dagger . \quad (43)$$

The replacement of $\mathbf{x}$ by $\mathbf{y}/\sqrt{\Delta\tau}$ introduces a nontrivial $\Delta\tau$ dependence into the one-electron potentials $\underline{V}_l$, and results in the factor $\frac{1}{2}$ multiplying it in the above expression. (If the factor were unity, the matrix elements above between the time slice eigenstates $\boldsymbol{\psi}_{jl}$ would be diagonal.) We recall from our discussion of the forces that equations of the same form as Eqs. (37) and (43) can be derived by considering the averages of one-electron operators carried out internally over the time slices. This enables us to rewrite the summand in Eq. (41) as a more physically transparent expression analogous to Eq. (23),

$$\left\langle \Phi_l^< \left| \frac{\partial \hat{U}_l}{\partial(\Delta\tau)} \right| \Phi_l^> \right\rangle = \mathrm{Det}_{ij}[S_{ij}^{\uparrow l}]\mathrm{Det}_{ij}[S_{ij}^{\downarrow l}] \left[\langle T\rangle^{\uparrow l} + \langle T\rangle^{\downarrow l} + \tfrac{1}{2}\sqrt{U_{\mathrm{Hub}}/\Delta\tau} \sum_{k=1}^{N} y_{kl}(\langle n_k\rangle^{\uparrow l} - \langle n_k\rangle^{\downarrow l}) \right], \quad (44)$$

where it must be remembered that the averages are in the sense just discussed, and not expectation values of operators between time slices. Aside from this distinction, we see that instead of evaluating the electron-electron interaction directly as in Eq. (23), we are evaluating the interaction of each electron density with the auxiliary field.

We briefly considered $E_4^*(\beta)$ early in our investigations, and discarded it because it has extremely bad $\Delta\tau$ convergence properties compared to any of our other expressions, as well as worse statistical properties. We were forced to reconsider it only after our numerical results indicated inconsistencies between our other measures and Eq. (40).

For the sake of completeness, we point out that had we gone through the above derivation without replacing $\mathbf{x}$ by $\mathbf{y}/\sqrt{\Delta\tau}$, we would have obtained yet another energy expression E_5. In this case, the $\frac{1}{2}$ factor would be replaced by unity in the matrix element in Eq. (43), and an additional $\frac{1}{2}\mathbf{x}^t * \mathbf{x}$ interaction term would appear. We anticipate that this measure would have even worse properties than E_4, and have not investigated it. Obviously, we could also have introduced Trotter approximation versions of E_4 and E_5.

III. RESULTS

We developed and tested the programs implementing the calculations already discussed using several two-dimensional Hubbard models for which exact diagonalization reference results were available. For the comparative study of the convergence of the several ground-state energy measurement strategies, we selected two systems that have a significant sign problem so we could simultaneously study the E^*-E energy difference. These are the simplest possible system, the 2×2 with three electrons, and the one with the worst sign behavior[13] we know of for which exact diagonalization results are available,[20] the 4×4 with 14 electrons. We found the convergence behavior of half-filled systems and other systems off half-filling to be similar in more limited studies. We chose $t=1$ and $U=4$ in both cases since the $4\times4/14$ diagonalization results were only available for these values.[20]

For both cases, we used a commensurate spin-density-wave-type state for $|\Phi_0\rangle$, generated by diagonalizing the Hamiltonian with an applied field of $+1$ on the A and -1 on the B sublattices. Results for the energy are identical using a paramagnetic state, and this choice served other purposes not discussed here. (For the half-filled case, such a choice is necessary in the projection method to complete the symmetry which forces positive signs.[14]) For these small systems, we found $E(\beta)$ to be independent of β for $\beta=5$, 10 and 20, and we chose $\beta=10$ for the $\Delta\tau$ convergence tests. We typically ran 500–1000 Monte Carlo steps for equilibrium, followed by 5000 measurement steps. The molecular dynamics run for each

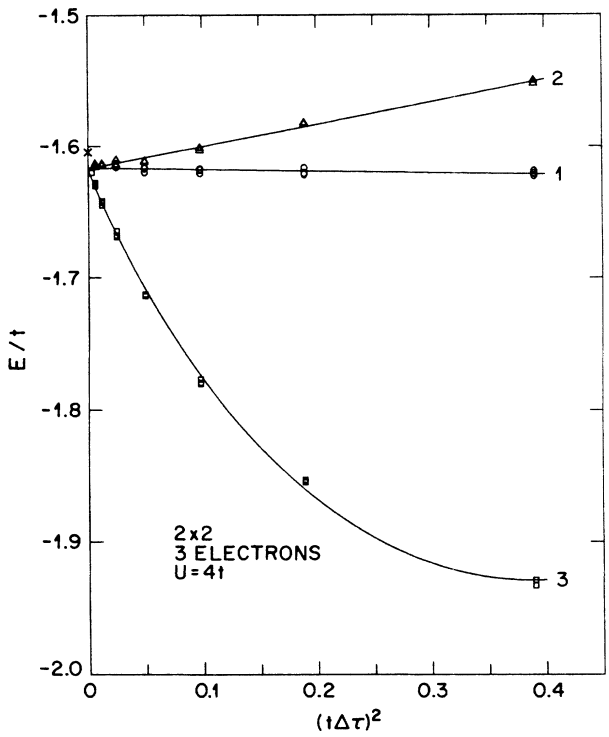

FIG. 1. Convergence of the absolute-value energy measures E^* with $(\Delta\tau)^2$ for the three-electron 2×2 Hubbard model. The lines are least-squares fits, and the differences among measures one, two, and three are discussed in the text. The "X" on the ordinate axis is our exact diagonalization result.

MC step used $n_{\mathrm{MD}}\Delta t\sim2.5$ as already discussed. We used blocks of 500 MC steps to determine our statistical error. While we have some evidence of longer correlation times, the E^* errors estimated this way (~0.001 to 0.002) are consistent with the run to run differences.

In Fig. 1, we plot $E^*(10)$ per site versus $\Delta\tau^2$ for the $2\times2/3$ case and our three approaches to energy measurement. The estimated statistical error just cited is comparable to the symbol size, and we performed three to four independent runs for most $\Delta\tau$ values. For the Trotter methods two and three, many of the measurements were made simultaneously on the same samples. The exact propagation operator, method one, of course, had different auxiliary-field dynamics and different state evolution, so all these runs were independent of the Trotter runs. The continuous curves are simple least-squares fits to the form

$$E(\Delta\tau)=E_i+b\Delta\tau^2+c\Delta\tau^3 . \tag{45}$$

The parameters and their estimated uncertainties are given in Table I. For E_1^* and E_2^*, parameter c was zero within its uncertainty, so the fit was redone with two terms only. We point out that all methods converge to the same value of E_i with high accuracy, which is an important consistency test. Striking differences in the $\Delta\tau$ convergence rates are immediately apparent from the figure, however. From the b parameters in the Table, it is clear that we achieve an order of magnitude improvement in going from method three to method two, and another order of magnitude in going to method one. The only quantitative result on convergence we have found in the literature is quoted by White *et al.* to be $b=-1.95\pm0.4$, with the uncertainty covering a wide range of Hubbard model parameters.[11] This is consistent with our method three result. We note that b is negative for method three, but positive for one and two. The X on the ordinate axis in Fig. 1 corresponds to the exact ground-state energy of the $2\times2/3$ model, which we obtained by direct diagonalization. It is clear that our E^* measures converge to a lower energy.

In Fig. 2, we show the corresponding results for the signed average E taken from the same runs. We found that the sign changes were quite accurately Poisson distributed over our Monte Carlo runs for this model, with an average of ~80 MC steps between changes. Our runs of 5000 steps therefore sampled both positive and negative regions with appropriate weights, and our sample should not have introduced any systematic error. As the scatter in the points indicates, the statistics are much worse, and we cannot reliably estimate the statistical error by other means. As seen from the figure and from

TABLE I. Ground-state energies per site for two Hubbard models with $t=1$ and $U=4$, and $\Delta\tau$ convergence parameters, based on least-squares fit to Eqs. (45) and (46). Standard deviations are given in parentheses. (Exact diagonalization energies are not fit.)

Model	Measure	E_i	a	b	c
$2\times2/3$	diag	-1.60463			
	1^*	$-1.6167(6)$		$-0.011(3)$	
	2^*	$-1.6170(11)$		0.16(3)	
	3^*	$-1.6184(17)$		$-2.40(5)$	2.56(8)
	4^*	$-1.6447(62)$	2.4(1)		
	1	$-1.6066(23)$		0.03(1)	
	2	$-1.6075(45)$		0.16(3)	
	3	$-1.6063(89)$		$-2.2(2)$	2.3(4)
$4\times4/14$	diag[a]	-0.9840			
	1^*	$-0.9951(14)$		0.013(8)	
	2^*	$-0.9962(15)$		0.161(7)	
	3^*	$-0.9935(37)$		$-1.87(8)$	1.6(1)

[a]Reference 20.

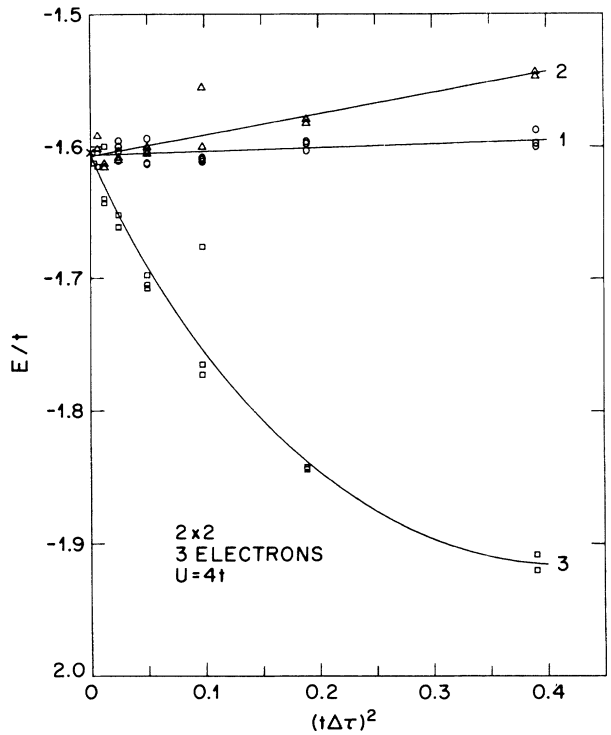

FIG. 2. Convergence of the signed energy measures, as in Fig. 1.

Table I, the convergence rates of the three methods are similar to those for the E^* calculation. All three converge to the exact diagonalization result, well within the uncertainty of the fit. This is further evidence for the correctness of our codes and our procedures, and for a small but significant difference between E^* and E. Using E_1^* and the exact diagonalization value to obtain our most accurate estimate, we obtain $\Delta E_1 = 0.012$. This implies, as discussed above and by Loh *et al.*,[13] that the sign $\langle s(\beta) \rangle$ vanishes exponentially with β.

To quantitatively test the relation between ΔE and the decay of $\langle s(\beta) \rangle$, we carried out an additional series of runs at a series of β values with $\Delta\tau$ fixed at 0.5. These were carried out using the exact propagators, which apparently give accurate results even for such a large value. To obtain accurate statistics, much larger samples were required, and we used runs of 50 000 and 100 000 Monte Carlo steps. These results are shown by the circles of Fig. 3. To confirm the $\Delta\tau$ independence of the $\langle s(\beta) \rangle$ decay, three additional calculations were carried out for $\Delta\tau = 0.25$, and these results are shown by triangles in Fig. 3. The error bars shown with these results are based on blocks of 5000 steps. A least-squares fit to all the points shown yielded $\ln\langle s(\beta) \rangle = 0.4679 - 0.1516\beta$. Fits dropping the $\beta = 4$ point yield essentially the same values, indicating the saturation of exponential behavior takes place quite abruptly at small β. Coverting the slope to a per site energy yields $\Delta E_s = 0.038$, where we have added the subscript s to indicate the source of this value.

Motivated by the discrepancy between our directly calculated ΔE and that derived from the sign decay, we returned to a careful analysis of the apparently violated Eq.

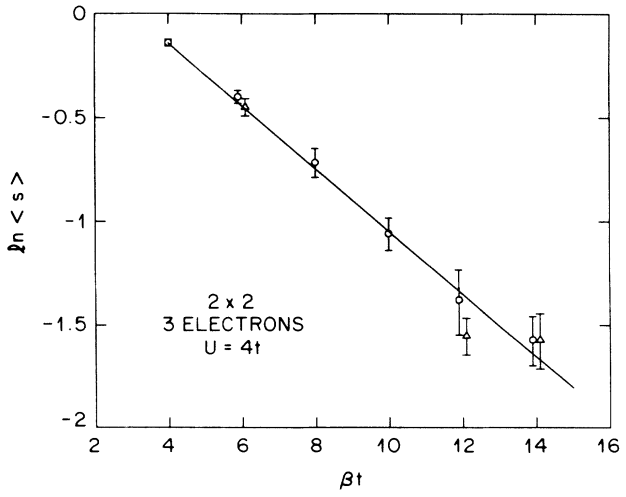

FIG. 3. Decay of the average value of the sign with β for the three-electron 2×2 Hubbard model.

(40), and to a calculation of E_4^* for this system. The $\Delta\tau$ convergence of E_4^* appears to be linear rather than quadratic, for reasons we do not understand. Even with the exact time-slice propagators, extremely small $\Delta\tau$ values are required to obtain a useful extrapolation to the limit. The results of a series of calculations at $\beta = 10$ using 5 000 Monte Carlo steps are shown in Fig. 4. The least-squares fit to

$$E(\Delta\tau) = E_4 + a\Delta\tau \tag{46}$$

is shown in the figure, and the parameters are given in Table I. We find $\Delta E_4 = 0.040(6)$, consistent with ΔE_s.

Figure 5 shows our results for the 4×4/14 model. As is apparent visually and from the fit parameters in Table I, the rates of $\Delta\tau$ convergence for E_1^*, E_2^*, and E_3^* are similar to those we found for the smaller model. The $(\Delta\tau)^2$ coefficient for method three, $-1.87(8)$, is even closer to the value of -1.95 given by White *et al.*[11] All

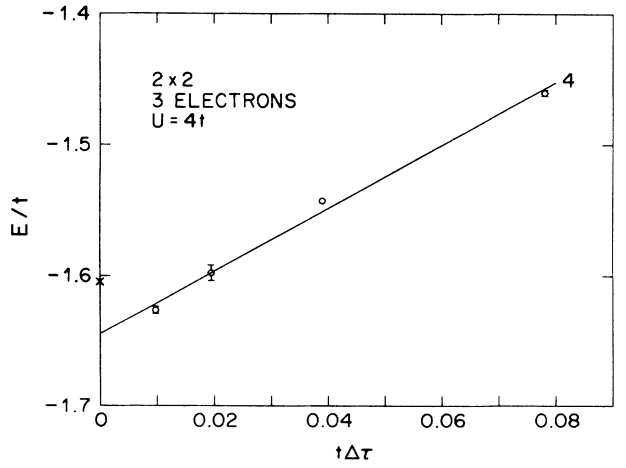

FIG. 4. Convergence of absolute-value energy measure four with $\Delta\tau$ for the three electron 2×2 Hubbard model. Note the different abscissa scale compared to Figs. 2 and 3.

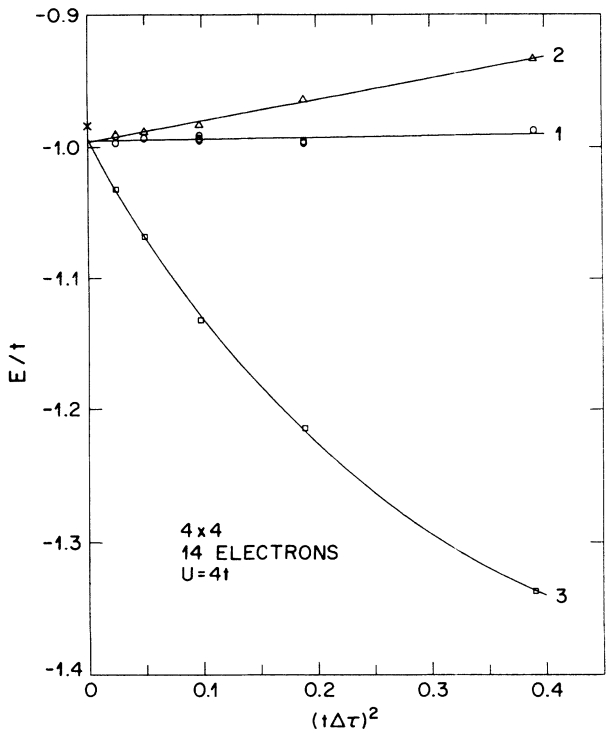

FIG. 5. Convergence of the absolute-value energy measures one through three for the 14 electron 4×4 Hubbard model. The "X" on the ordinate axis is the exact diagonalization result from Ref. 20.

three methods converge to the same E^* value, and $\Delta E_1 = 0.011$ is nonzero well outside our uncertainty.

For the 4×4/14 model, we also find sign changes Poisson distributed over the Monte Carlo runs, with a mean spacing of ~50 steps. We found that $\langle s(\beta=10)\rangle$ is consistent with 0 within our statistics, and we did not attempt to calculate the decay of $\langle s(\beta)\rangle$ ourselves. For this case, however, a plot of $\ln\langle s\rangle$ versus β is given by Loh *et al.*[13] Based on their straight line fit, we infer an energy correction $\Delta E_s = 0.044$ per site. The inconsistent ΔE_1 and ΔE_s values for the 4×4/14 model are quantitatively nearly identical to those for the 2×2/3 model.

IV. DISCUSSION

One of our key results is the improvement in $\Delta\tau$ convergence over the "industry standard" that results from switching from method three to method two. Both use the Trotter approximation,[16] but differ in where the energy measurement is performed. In both cases the full many-body Hamiltonian is used for the energy measurement. We could imagine we arrived at method three in the following manner: Starting from the Eq. (39), we introduce the discrete $\tau=\beta/P$ mesh and apply the Trotter approximation to each time slice propagator while still in many-body form. Then we carry out the $d/d\beta$ derivative to bring down the kinetic-energy and Hubbard interaction operators for energy measurement. Then we carry out the Hubbard-Stratonovich transformation. For method two, we followed, in principal, an alternative sequence. Starting from Eq. (5), we replaced the τ integral by a discrete sum that just happened to coincide with our post Hubbard-Stratonovich transformation τ mesh. Next, we carry out the transformation in the numerator and denominator of Eq. (5), considering our time ordering to be based on infinitesimal $d\tau$ increments. Next, we replace $\mathbf{x}(\tau)$ by the stepwise approximation to the continuous τ function. Finally, we carry out the Trotter breakup on the individual time slices.

If one takes the point of view that the finite $\Delta\tau$ Trotter approximation must be introduced to justify the Hubbard-Stratonovich transformation, then both methods should be equally good, or rather equally bad for large $\Delta\tau$. The symmetric Trotter approximation is accurate to order $(\Delta\tau)^3$, and the integrated error from 0 to β should then be of order $(\Delta\tau)^2$. Our results do not suggest the order of performing the Hubbard-Stratonovich transformation and the Trotter approximation change the power with which the energy converges, just the prefactor.

We can arrive at our method one by following the development of two as already outlined, and simply leaving out the Trotter step. In this case we can offer a justification for the improvement. We have studied the Fourier spectra of the auxiliary fields as a function of τ. There are as many Fourier components as there are time slices, so smaller $\Delta\tau$ in effect introduces higher frequencies. One can demonstrate in model calculations that very high frequency fluctuations in the fields have very little effect on the states being propagated.[14] Our measured spectra confirm the inverse effect, namely that the high components are strictly Gaussian in their distribution, showing that the determinant of states reacts back on them with very weak forces. The spectrum cuts off rapidly, with $\omega = 2\pi n/\beta \sim U$ as the apparent characteristic frequency (we have not checked this dependence systematically). The amplitudes of the lower components do not vary much when higher components are introduced through smaller $\Delta\tau$. All this supports the idea that the higher components are not very important to the development of the projected many-body state.

However, a small $\Delta\tau$ is important to obtain accuracy from the Trotter approximation. Early in our explorations of this method, we performed a test in which we set up a spatially random but τ-independent set of fields with Gaussian distributed amplitudes. We used the eigenstate of this field configuration as $|\Psi_0\rangle$, and measured the error in the one-electron energy produced by propagating this state with a sequence of (identical) Trotter $\underline{U}_l$'s. While quantitative comparison between this test and our many-body results is not possible, the $(\Delta\tau)^2$ dependence of the errors was of similar magnitude.

Our conclusion from these observations is that the Trotter formula and correct treatment of the many-body effects are not inexorably linked. Applying the approximation later is better, and never applying it is best. Most of the computational work entailed by small $\Delta\tau$ is spent doing the trivial part of the problem, the one-electron part, with sufficient accuracy. When these methods are applied to real systems, this will be especially important

because of the large ion and quasistatic screening potentials, which are likely to completely dominate the dynamically fluctuating potentials.

The other issue we have addressed is the energy difference ΔE resulting from averaging the energy contributions from positive and negative sign regions rather than taking their difference. For both cases, we found small but statistically significant ΔE values. This is in contrast to results reported in Refs. 8 and 20, where ΔE was found to be zero within statistical error for the cases reported. The energy E^* for the $4\times4/14$ case is reported to be $-0.988(3)$ in Ref. 20, which is in sufficient disagreement with our result of $-0.9951(14)$ to lead to different conclusions about ΔE. We cannot assess the cause of the discrepancy, if it is real and not statistical, on the basis of the limited information given, but there are several possibilities. As discussed in Refs. 7 and 8, the Trieste group uses Langevin dynamics to sample the auxiliary fields. It is not clear whether their $4\times4/14$ results are calculated using the $|\langle\Phi_0|\hat{U}|\Phi_0\rangle|$ "action" or their proposed $(\langle\Phi_0|\hat{U}^\dagger\hat{U}|\Phi_0\rangle)^{1/2}$ action.[8] We have found that both give similar results, and that the second, while positive definite by construction, has near zeroes that are computationally equivalent to the true zeroes of the first. The Langevin method must either fail to penetrate these barriers if used with small enough time step, or give errors throughout if used with a time step designed to facilitate barrier penetration. The other possible cause of the disagreement is $\Delta\tau$ convergence, which is not discussed.[20]

The comparisons among ΔE_1, ΔE_4, and ΔE_s are the least well understood part of our result. As this research developed, we first became aware of a discrepancy between our ΔE_1 for the $4\times4/14$ model and ΔE_s after receiving the unpublished $\langle s(\beta)\rangle$ result of Loh *et al.*[13] These authors use rather different methods (e.g., Hirsch's discrete Hubbard-Stratonovich transformation[12]), and do not give any energy results. While we first considered that the discrepancy might arise from the methodological differences, we found this rationalization unconvincing. This led us to confirm the similar discrepancy for the $2\times2/3$ model within a fully consistent computational context, and finally to the "resolution" of the discrepancy through our E_4^* study.

Schematically, all the energies have the form indicated in Eq. (28), a functional integral of a weight factor GD times an energy functional, appropriately normalized. If we confine our comparison to measures one and four, where the time-slice propagator is treated exactly, only the energy functionals $E_1(\mathbf{x})$ and $E_4(\mathbf{x})$ differ. When they are averaged keeping the sign of D, they must both converge to the same number at small $\Delta\tau$. Yet when they are averaged using $|D|$, they give different results. The E_1^* average is relatively easy to compute and is closer to the exact result. Yet we apparently have no theory for correcting it to the exact result. The E_4^* average is much more difficult to compute and farther from the exact result, but we do have a theoretical expression for the correction.

For application to realistic many-electron systems or to large models, the signed averages will probably never be practical to compute with the required accuracy. The observation by Sorella *et al.*[7,8] that the absolute value measure gives energies that are very close to exact values is a significant advance. To realize its full advantage, however, it will be necessary to develop a theory and a computational approach to systematically estimate the correction. The present results indicate that this issue is more subtle than previously suspected.

ACKNOWLEDGMENTS

The authors wish to thank S. R. White, S. Baroni, and R. Car for valuable discussions, and for making Refs. 8, 11, 13, and 20 available prior to publication.

APPENDIX

In this appendix, we give simple derivations of several results that were used in the text, and of the connection between the projection method[6,7] and the grand canonical ensemble method.[11]

Suppose we have an arbitrary $N\times N$ matrix $\underline{A}$ in the one-electron basis, and form the bilinear Fock-space operator

$$\hat{A}=\mathbf{c}^\dagger*\underline{A}*\mathbf{c}\ . \tag{A1}$$

Now using the elementary anticommutation relations of the electron creation and annihilation operators, we can show that

$$\hat{A}\mathbf{c}^\dagger=\mathbf{c}^\dagger*(\hat{A}\underline{I}+\underline{A})\ , \tag{A2}$$

where $\underline{I}$ is the $N\times N$ unit matrix. Note that this relation combines in a compact notation the matrices $\underline{A}$ and $\underline{I}$, whose elements are c numbers in Fock space, with $\hat{A}$, which is an operator in Fock space, but is a scalar relative to the $N\times N$ matrices. From Eq. (A2), it follows that

$$\hat{A}^n\mathbf{c}^\dagger=\mathbf{c}^\dagger*(\hat{A}\underline{I}+\underline{A})^n\ , \tag{A3}$$

and

$$e^{\hat{A}}\mathbf{c}^\dagger=\mathbf{c}^\dagger*e^{(\hat{A}\underline{I}+\underline{A})}\ . \tag{A4}$$

Now, from the properties of $\hat{A}\underline{I}$ and $\underline{A}$ discussed above they commute both as matrices and as Fock-space operators, so

$$e^{\hat{A}}\mathbf{c}^\dagger=\mathbf{c}^\dagger*e^{\underline{A}}e^{\hat{A}}\ . \tag{A5}$$

This result allows $e^{-\Delta\tau\hat{H}}$ to "walk through" the string of creation operators defining $|\Phi\rangle$, replacing each one-electron wave function $\boldsymbol{\phi}$ by the propagated wave function $e^{-\Delta\tau\underline{H}}*\boldsymbol{\phi}$ in Eq. (14).

It can be shown from the elementary commutation relations of the electron operators that the overlap of two M-electron Slater determinants, defined as in Eq. (2), with different sets of one-electron wave functions $\boldsymbol{\psi}$ and $\boldsymbol{\phi}$ is given by

$$\langle\Psi|\Phi\rangle=\mathrm{Det}_{ij}[S_{ij}]\ , \tag{A6}$$

where the $M\times M$ overlap matrix S_{ij} is simply

$$S_{ij}=\boldsymbol{\psi}_i^\dagger*\boldsymbol{\phi}_j\ . \tag{A7}$$

The matrix element of the bilinear operator $\hat{A}$ is easily expressed, using Eq. (A2), as

$$\langle\Psi|\hat{A}|\Phi\rangle=\sum_{k=1}^{M}\mathrm{Det}_{ij}\{\psi_i^\dagger*[(1-\delta_{jk})\underline{I}+\delta_{jk}\underline{A}]*\phi_j\}\ .$$

If we expand each determinant in the sum in terms of the cofactors (signed minors) s of column k, which are the same as the cofactors of S_{ij}, we obtain

$$\langle\Psi|\hat{A}|\Phi\rangle=\sum_{l,k=1}^{M}(\psi_l^\dagger*\underline{A}*\phi_k)s_{lk}\ , \tag{A8}$$

which, using Cramer's rule, can be expressed in terms of the inverse of S,

$$\langle\Psi|\hat{A}|\Phi\rangle=\mathrm{Det}[S]\sum_{k,l=1}^{M}(S^{-1})_{kl}(\psi_l^\dagger*\underline{A}*\phi_k)\ . \tag{A9}$$

This result was used in Eqs. (23)–(25) and (27).

In the finite-temperature auxiliary-field method based on the grand canonical ensemble, the partition function is expressed as the Gaussian-weighted integral of the $N\times N$ determinant[11]

$$D=\mathrm{Det}_{ij}[(\underline{I}+\underline{U})_{ij}]\ , \tag{A10}$$

where $\underline{U}$ is the ordered product of the time-slice propagation matrices,

$$\underline{U}=\underline{U}_P*\cdots*\underline{U}_2*\underline{U}_1\ . \tag{A11}$$

Note that in general $\underline{U}$ cannot be expressed as the exponential of a Hermitian matrix, even though the individual time-slice propagators have this form. Equation (A10) is derived using a combination of Grassmann algebra and Green's function techniques,[21] and its relation to the projection technique, to its interpretation in terms of propagating individual Slater determinants, and to the $M\times M$ determinant $\langle\Phi|\underline{U}|\Phi\rangle$ is not transparent.

Any complete orthonormal basis set $\{\phi_i\}$ forms a unitary matrix, and since the determinant of a matrix is invariant under unitary transformation,

$$D=\mathrm{Det}_{ij}[\phi_i^\dagger*(\underline{I}+\underline{U})*\phi_j]\ . \tag{A12}$$

This expression can be expanded as a sum of 2^N determinants, since each ϕ_j can be acted upon by either $\underline{I}$ or $\underline{U}$,

$$D=\sum_{M=0}^{N}\sum_{\{C\}}\mathrm{Det}_{ij}[D_{ij}^{MC}]\ , \tag{A13}$$

where $\{C\}$ is the set of ${}^NC_M$ combinations of M integers drawn from the set $\{1,\ldots,N\}$, and

$$D_{ij}^{MC}=\begin{cases}\phi_i^\dagger*\underline{U}*\phi_j, & j\in C\ ,\\ \delta_{ij}, & j\notin C\ .\end{cases} \tag{A14}$$

The Kronecker delta occurs because ϕ's are orthonormal. In effect, each wave function acted upon by $\underline{U}$ is propagated, while each acted upon only by $\underline{I}$ is passively carried to β. Now for each element in the sum, we can reorder the set of ϕ's to put all the propagating wave functions first, since this subjects the rows and columns of the determinant to the same permutation, and does not change its value. The rearranged matrix is

$$\tilde{D}_{ij}^{\mathrm{MC}}=\begin{cases}\phi_i'^\dagger*\underline{U}*\phi_j', & j=1,M\ ,\\ \delta_{ij}, & j=M+1,N\ ,\end{cases} \tag{A15}$$

where the primes simply denote the new indexing of the wave functions. We now observe that all the elements of $\tilde{D}$ in columns $1,M$ and rows $M+1,N$ can be zeroed by Schmidt orthogonalization to columns $M+1,N$, so the matrix reduces to our familiar $M\times M$ block in the upper left, ones on the remaining diagonal, and zeros elsewhere.

We have thus shown that each term of D is equal to some $\langle\Phi|\underline{U}|\Phi\rangle$ of the projection method, and that we must sum over the set of all possible Slater determinant trial functions to get the partition function at finite β. Of course a chemical potential must be included, and at large β those $|\Phi\rangle$ with the most probable M and maximum overlap with the ground state will dominate the sum.

[1]W. L. McMillan, Phys. Rev. A **138**, 442 (1965).

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

[3]S. B. Fahy, X. W. Wang, and S. G. Louie, Phys. Rev. Lett. **61**, 1631 (1988).

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

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

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

[7]S. Sorella, E. Tosatti, S. Baroni, R. Car, and M. Parrinello, Int. J. Mod. Phys. B **1**, 993 (1988).

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

[9]R. L. Stratonovich, Dokl. Akad. Nauk. SSSR **115**, 1097 (1957) [Sov. Phys.—Dokl. **2**, 416 (1957)].

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

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

[12]J. E. Hirsch, Phys. Rev. B **28**, 4059 (1983).

[13]E. Y. Loh, Jr., J. E. Gubernatis, R. T. Scalettar, S. R. White, D. J. Scalapino, and R. L. Sugar (unpublished).

[14]S. B. Fahy and D. R. Hamann (unpublished).

[15]D. R. Hamann, Phys. Rev. B **2**, 1373 (1970).

[16]H. F. Trotter, Proc. Am. Math. Soc. **10**, 545 (1959); M. Suzuki, Commun. Math. Phys. **51**, 183 (1976).

[17]Brief experience with the nonsymmetrized version of Eq. (26) gave considerably worse results, and it was not pursued.

[18]S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, Phys. Lett. B **195**, 216 (1987). See also R. T. Scalettar, D. J. Scalapino, and R. L. Sugar, Phys. Rev. B **34**, 7911 (1986).

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

[20]A. Parola, S. Sorella, S. Baroni, R. Car, M. Parrinello, and E. Tosatti (unpublished).

[21]R. Blankenbecler, D. J. Scalapino, and R. L. Sugar, Phys. Rev. D **24**, 2278 (1981).