Global Monte Carlo algorithms for many-fermion systems

Michael Creutz

Physics Department, Brookhaven National Laboratory, Upton, New York 11973

I discuss algorithms for simulating many-fermion systems via global updatings of auxiliary fields followed by an accept-reject stage which eliminates finite-step-size errors. When the system size is larger than the correlation length, these procedures should require computer time growing only slightly faster than linearly with the system volume V. A corrected Langevin scheme should asymptotically display a $V^{4/3}$ behavior, while the hybrid Monte Carlo scheme can behave as $V^{5/4}$. I present some tests of the latter algorithm on a simple model of interacting electrons on a two-dimensional lattice.

I. INTRODUCTION

While Monte Carlo methods have become a prime tool in the study of quantum field theories, the major successes have been for systems involving only bosonic fields. Although many recent and ongoing calculations are obtaining interesting results for systems containing dynamical fermions, this is only at the expense of great strain on computational resources. While the situation will inevitably improve with better technology, this is an area where substantial further gains may arise through algorithm development

In addition to the usual approximations involved in formulating a continuum theory on a finite lattice and in a finite volume, most of the fermionic algorithms used in practice make an additional extrapolation in a step-size parameter. Here I include the original pseudofermionic[1] technique as well as the microcanonical,[2] Langevin,[3,4] and hybrid[5] approaches, which involve discretization of a differential evolution. Although systematic effects associated with this extrapolation are occasionally studied, the severe computational demands constrain the abilities to make as many checks as might be desired. While these difficulties will lessen with improved computational facilities, an efficient scheme for fermionic updating without this additional approximation would be welcome.

Unfortunately, the known "exact" algorithms tend to require substantially more computer time than those making finite-step approximations. In addition, the required computer time increases rapidly with the system size; the prototype exact algorithm presented by Weingarten and Petcher[6] requires a time growing as the volume squared of the system.

For the exact fermionic algorithms, this growth of computer time with volume is not universal, and can be reduced by a suitably biased selection of trial changes. Recently it has been proposed[7–12] to modify the Langevin and microcanonical updating schemes into exact algorithms by combining them with a global Metropolis *et al.*[13] acceptance step. The purpose of this paper is to study these methods further. One of the main points is that these schemes potentially require computer time which grows only slightly more rapidly with the system

volume than the linear behavior of bosonic algorithms.

Section II introduces the formal problem and the Weingarten-Petcher[6] formalism underlying the later algorithms. Section III discusses global accept-reject procedures with biases in the trial selection process. Here I argue for the favorable volume dependence of the methods presented in Refs. 7 and 9. Section IV introduces the many-electron model on which I test variations of these algorithms. Section V contains concluding remarks and speculations. The Appendix gives some useful connections between fermionic operators and integrals over anticommuting fields.

II. FERMIONS AND MONTE CARLO SIMULATION

I begin my discussion with a generic path integral or partition function for a system of fermionic fields ψ and $\bar{\psi}$ quadratically coupled to a bosonic field A:

$$Z = \int (d\psi \, d\bar{\psi})(dA)\exp[-S_0(A)-\bar{\psi}M(A)\psi] . \quad (1)$$

Here S_0 combines the free action and self-couplings for the field A, and the interaction with the fermions is through the matrix $M(A)$. I suppress the matrix and vector indices on M and the fermionic fields. These indices have a range proportional to the system volume. In addition to the gauge theory of quarks and gluons, numerous many-fermion quantum systems can be cast in this form.

Not knowing how to deal directly with a large number of anticommuting variables, I immediately integrate out these fields to obtain[14]

$$Z = \int (dA)e^{-S_0(A)}\det[M(A)] . \quad (2)$$

For a Monte Carlo simulation of this system, I need to find configurations of the field A with a probability distribution

$$P_{\text{eq}}(A) \propto e^{-S_0(A)}\det[M(A)] . \quad (3)$$

I will assume at the outset that $\det[M(A)]$ is a nonnegative quantity so that Eq. (2) makes sense as a probability distribution. I also will assume that M is a real matrix. This is mainly for simplicity in the following discus-

sions, although certain complications arise with complex M.

Weingarten and Petcher[6] presented a conceptually simple framework for obtaining this probability distribution. Introducing an auxiliary real field ϕ, they observe that the partition function in Eq. (2) is proportional to the integral

$$Z \propto \int (dA)(d\phi)\exp\{-S_0-[M^{-1}(A)\phi]^2/2\} . \quad (4)$$

The integration over ϕ gives the desired factor of $\det M$. Thus, the problem reduces to finding configurations of fields A and ϕ with probability distribution

$$P_{eq}(A,\phi) \propto \exp\{-S_0-[M^{-1}(A)\phi]^2/2\} . \quad (5)$$

If M is a complex matrix, then one should use a complex ϕ and take an absolute value squared for the second term in the exponent of Eq. (5). However, this would give the square of the determinant of M in the probability distribution for A. Effectively this would double the number of fermionic species. Various schemes have been proposed to reduce this extra doubling, which I will ignore here. The model discussed later in this paper avoids this problem by using a real matrix M.

Reference 3 presented an efficient scheme for updating the field ϕ while holding A fixed. First, generate a random real vector χ with a Gaussian probability distribution

$$P(\chi) \propto e^{-\chi^2/2} . \quad (6)$$

Then construct

$$\phi = M\chi . \quad (7)$$

This change of variables gives ϕ with the desired probability. The Jacobian of the transformation is irrelevant because A is temporarily being held fixed.

Updating the A field remains the slow part of any algorithm based on Eq. (5) because of the appearance of $M^{-1}(A)$. As this inverse appears only applied to the single vector ϕ, a linear equation-solving algorithm, i.e., conjugate gradient, can be used to calculate the action

$$S = S_0(A)+[M^{-1}(A)\phi]^2/2 . \quad (8)$$

For an updating scheme like that of Metropolis et al.[13] this quantity must be calculated for both current and trial values for the field A. If this is done every time a single degree of freedom is updated, such an algorithm requires computer time growing as the volume of the system squared. One factor of volume comes from the sweep over the system variables, and another from the optimistic assumption that the conjugate-gradient algorithm will adequately converge in a fixed number of iterations, each of which requires a time growing linearly with the volume.

This time growth as the volume squared is common to many exact algorithms, and is considerably better than the naive direct calculation of the determinant of M, which would give a fourth power of the volume behavior. In this and the following discussion of volume dependences, I assume that all couplings are being held fixed and ignore any additional slowing from critical behavior. In particular, this means that one is not at a critical point

and all correlation lengths are finite.

Because of the severe computational requirements with this algorithm, most current simulations make an additional approximation involving a small step size. The hope is to avoid calculating inverses after each change by assuming that small changes will allow but a single conjugate-gradient inversion per sweep of the entire lattice. At the end of any calculation one should extrapolate observables to a vanishing value for the size of the trial changes. This forms the basis of most of the popular fermionic methods.[1-5] For a review emphasizing the similarities of the various algorithms see Ref. 15.

In numerical simulations one already approximates continuum physics by using a finite lattice spacing and working in a finite volume. To introduce an additional approximation involving a new extrapolation is rather unaesthetic. This motivates further search for so-called "exact" algorithms which do not introduce finite-step-size errors. For the remainder of this paper I discuss a class of such algorithms involving accept-reject steps on global sets of variables.

III. GLOBAL ACCEPT-REJECT ALGORITHMS

The difficulty with calculating the quantity $M^{-1}(A)\phi$ appearing in the exact action of Eq. (8) strongly encourages algorithms which perform this computation as rarely as possible. Indeed, one of the prime motivations for the approximate algorithms is to do this inversion only once per sweep of the lattice variables. There are, however, exact algorithms which also perform this inversion only once per sweep. In particular, one can perform a Metropolis et al. type accept-reject step to restore exact detailed balance only after making trial changes over the whole system. This forms the basis for the algorithms discussed in Refs. 7–12.

The concept of applying an acceptance condition to a global change has the danger that one will have large increases in the action if many variables are changed and thus, final acceptances may be unfeasibly small. Indeed, an arbitrary change in all variables will increase the action by an amount proportional to the system volume, leading to an exponential suppression of the acceptance. To counteract this, one must reduce the step size as the volume increases. By appropriately biasing the trial changes, one may be able to keep the acceptance at a reasonable level for step sizes which still allow a practical rate of exploration of new configurations.

To make this more precise, consider a trial change of a single variable A to

$$A' = A +p\delta+F(A)\delta^2 . \quad (9)$$

Here δ is an adjustable step-size parameter introduced for bookkeeping purposes. The "momentum" variable p represents a random noise, which for convenience I take to be in a Gaussian distribution:

$$P(p) \propto e^{-p^2/2} . \quad (10)$$

The function $F(A)$ represents a driving force or bias in the trial selection procedure and is for the moment arbitrary.

The Metropolis *et al.*[13] scheme accepts trial changes with a conditional probability chosen to maintain detailed balance when applied to an equilibrium ensemble. With an unbiased trial change this acceptance is determined entirely by the exponentiated action change. Here, however, the force term in the selection procedure must be corrected for in the acceptance condition. I fully restore detailed balance by accepting the new value A' with probability

$$P_{\text{acc}} = \min[1, e^{H(p, A) - H(p', A')}] \ . \tag{11}$$

Here H is a classical "Hamiltonian"

$$H(p, A) = p^2/2 + S(A) \ . \tag{12}$$

In Eq. (11) I introduce p' as the negative of the reverse noise, i.e., the noise which would be required for the selection of A as the trial had A' been the initial value

$$p' = p + [F(A) + F(A')]\delta \ . \tag{13}$$

Note that H is precisely the Hamiltonian used in the microcanonical algorithm[2] to describe evolution in "simulation time." Because of this analogy, I refer to H as the classical energy.

Equations (9) and (13) considered together represent a discretization of a microcanonical step of A and p under the Hamiltonian in Eq. (12). The microcanonical time step is δ. Furthermore, the second-order terms are constructed to ensure that this particular mapping exactly preserves areas in phase space

$$dA \, dp = dA' \, dp' \ . \tag{14}$$

Note that the overall algorithm drives p and A towards equilibrium with the coupled probability

$$P(A, p) \propto e^{-S(A) - p^2/2} \ . \tag{15}$$

Some rather useful identities follow from considering expectation values over this distribution. Consider the partition function

$$Z = \int dA' dp' e^{-H'} = \int dA \, dp \, e^{-H} e^{H - H'} \ , \tag{16}$$

where H and H' denote $H(p, A)$ and $H(p', A')$, respectively. Dividing by Z, I find

$$\langle e^{H - H'} \rangle = 1 \ , \tag{17}$$

where the expectation value is over initial p and A distributed as in Eq. (15). By Jensen's inequality this immediately implies

$$\langle H' - H \rangle \geq 0 \tag{18}$$

with equality only possible if the algorithm exactly conserves energy. Below I will be considering small changes in the energy, in which case the expansion of Eq. (17) to second order,

$$\langle H' - H \rangle = \tfrac{1}{2} \langle (H' - H)^2 \rangle + O((H' - H)^3) \ , \tag{19}$$

will prove to be very useful.

To proceed I use the parameter δ for an expansion of the energy change. Doing some straightforward algebra, I find

$$H' - H = \left[p\delta + \frac{1}{2} \left[p^2 \frac{\partial}{\partial A} + 2F(A) \right] \delta^2 \right]$$

$$\times \left[\frac{\partial S(A)}{\partial A} + 2F(A) \right] + O(\delta^3) \ . \tag{20}$$

Note that the choice

$$F_L(A) = -\frac{1}{2} \frac{\partial S}{\partial A} \tag{21}$$

leads to an energy change which starts at third order in δ. Indeed, making this choice and ignoring the possibility of rejecting the trial change gives the usual Langevin algorithm,[3,4] where the parameter δ is the square root of the step size used for discretization.

First I consider not making the Langevin choice for the driving force. Inserting Eq. (20) into Eq. (19) gives

$$\langle H' - H \rangle_{A, p} = \frac{\delta^2}{2} \left\langle \left[\frac{\partial S(A)}{\partial A} + 2F(A) \right]^2 \right\rangle + O(\delta^4) \ . \tag{22}$$

Note that terms with odd powers of δ in the energy change expansion all involve odd powers of p and thus vanish in this average. If I now consider updating some large number V of variables together, the positive $O(\delta^2)$ quantities will coherently add and I expect to find a total energy change increasing linearly with V. By the central limit theorem, the fluctuations about this growth will become Gaussian. Thus, for large volumes, I expect to find

$$H' - H \simeq C\delta^2 V + B\rho\delta V^{1/2} \ , \tag{23}$$

where C and B are constants and ρ is a Gaussian random variable which I normalize such that its probability distribution is

$$P(\rho) \sim e^{-\rho^2/2} \ . \tag{24}$$

If Eq. (23) were exact, then Eq. (17) would relate C and B:

$$C = B^2/2 \ . \tag{25}$$

With this explicit form for the energy change, I can obtain the expected acceptance in the large-V limit

$$\langle P_{\text{acc}} \rangle = \langle \min[1, e^{H - H'}] \rangle$$

$$= \frac{2}{(\pi C V \delta^2)^{1/2}} e^{-C V \delta^2/4} \left[1 + O\left[\frac{1}{C V \delta^2} \right] \right] \ . \tag{26}$$

The calculations required to derive Eqs. (25) and (26), however, depend strongly on the tails of the distribution of the energy change and thus cannot be regarded as completely rigorous. For the following I will only assume that the expected acceptance is exponentially suppressed when $V\delta^2$ is large.

To avoid this exponential suppression and have a reasonable acceptance requires $\delta \sim V^{-1/2}$. However, a small value for δ raises the issue that the lattice will evolve only slowly away from its original configuration. More precisely, consider taking N sweeps over the lattice. As the motion of A it has both random and driven terms, the overall change in any given variable should go as

$$\Delta A = O(\delta\sqrt{N}) + O(\delta^2 N) = O(\sqrt{N/V}) + O(N/V) .$$

(27)

The final result is that the number of sweeps required to obtain a substantially new configuration should grow as V. If V is proportional to the system volume, then the overall algorithm requires time growing as volume squared, one factor of volume from the number of sweeps, and the other from the fact that each sweep takes time proportional to the volume.

For a bosonic simulations this growth would be a disaster. The conventional algorithms only grow as the system volume, and thus any gain involving updating many variables simultaneously would require some major additional cleverness. However, for exact fermionic algorithms we already are starting with a volume squared behavior, and thus there is no obvious major penalty in going to global updates. Indeed, it might be possible to gain something by a judicious choice of F which will reduce the coefficient of this growth.[15]

I now return to the Langevin choice of Eq. (21) for the driving force. To begin, consider again updating only a single variable. At first glance one might think that since the exact action is so difficult to calculate, the requisite derivative for this force would be intractable. Luckily, as everybody doing fermionic simulations with either Langevin or microcanonical methods knows, this is not the case. Considering the action in Eq. (8), the force takes the form

$$F_L(A) = -\frac{1}{2}\frac{\partial S_0}{\partial A} + \frac{1}{2}\left[M^{\dagger -1}M^{-1}\phi, \frac{\partial M}{\partial A}M^{-1}\phi\right] . \quad (28)$$

In addition to $M^{-1}\phi$ appearing in the action, calculation of the force requires knowledge of $(MM^{\dagger})^{-1}\phi$. This can also be found via whatever linear equation solver is being used. Indeed, if M is not Hermitian, standard inversion algorithms calculate this quantity anyway. Thus, the Langevin choice of the driving force does not present any major new obstacles.

To proceed I slightly generalize this force and take

$$F(A) = -\frac{1}{2}\frac{\partial S}{\partial A} + g(A)\delta^2 . \quad (29)$$

The $g\delta^2$ piece is included to allow higher-order improvements. Using this, I calculate the next term in the expansion for the energy change

$$H' - H = -\frac{\delta^3}{12}(S_3 p^3 - 3S_1 S_2 p - 24gp) + O(\delta^4) . \quad (30)$$

Here I use the notation

$$S_n = \frac{\partial^n S}{\partial A^n} . \quad (31)$$

Note that if S_3 is nonvanishing, i.e., if the theory is not harmonic, then no choice of $g(A)$ can make the $O(\delta^3)$ term in this equation vanish for all p. Thus, for any driving force whatsoever, a Metropolis *et al.*[13] correction to the Langevin evolution will always reject some changes if the step size is finite.

I am ultimately interested in applying this procedure to a group of variables simultaneously. Thus, as in the earlier discussion of unbiased changes, I am interested in any coherent addition of action changes which could give an exponential suppression of the final acceptance. Because the above $O(\delta^3)$ term contains only odd powers of p, it will vanish on the average. Indeed, the expectation of the energy change will only involve even powers of δ. The expectation for the energy change can again be most easily found using Eq. (19). With a little algebra and explicitly doing the average over p, I find

$$\langle H' - H \rangle = \frac{\delta^6}{96}\langle 2S_3^2 + 3(8g - S_3 + S_1 S_2)^2 \rangle + O(\delta^8) . \quad (32)$$

This can be written in many forms; this expression as the sum of two squares emphasizes positivity.

For nonharmonic interactions the quantity in Eq. (32) is positive for any g. Thus, one cannot use the freedom in redefining the force to push the expected energy change to higher order. This result does, however, suggest that an improved driving force is

$$F(A) = -\frac{1}{2}\frac{\partial S}{\partial A} + (S_3 - S_1 S_2)\delta^2/8 . \quad (33)$$

Note that this choice also eliminates the harmonic pieces in Eq. (30). Difficulties with calculating the higher derivatives of the action may preclude the utility of this form.

I now return to updating a large number V of independent variables simultaneously. The positive contributions indicated in Eq. (32) will add coherently. Similar arguments to those leading to Eq. (26) now give an expected acceptance falling as

$$P_{acc} \sim e^{-CV\delta^6} . \quad (34)$$

To have a reasonable acceptance requires only $\delta \sim V^{-1/6}$. This changes Eq. (24) to

$$\Delta A = O(\delta\sqrt{N}) + O(\delta^2 N)$$
$$= O(N^{1/2}/V^{1/6}) + O(N/V^{1/3}) . \quad (35)$$

The number of sweeps for an independent lattice grows as $V^{1/3}$ and the overall computer time for decorrelation increases as

$$T \sim V^{4/3} . \quad (36)$$

This behavior is only slightly worse than the linear growth of the pure bosonic theory.

This algorithm was proposed in Ref. 7 and tested further with somewhat discouraging results in Ref. 8. Reference 9 presents an even more promising variation on the above. Recapitulating on the above discussion, I constructed both the trial new A and the negative of the noise needed to return

$$A' = A + p\delta + F(A)\delta^2 , \quad (37a)$$

$$p' = p + [F(A) + F(A')]\delta . \quad (37b)$$

The scheme proposed in Ref. 9 is to iterate Eq. (37) several times before making the accept-reject decision.

This iterated map remains reversible and area preserving. The second-order terms in this equation make it equivalent to the leapfrog procedure with an initial half-step as used in Ref. 9.

The important point is that after each step the momentum remains the negative of that which would be required to reverse the motion and return exactly to the initial variables. Thus, a final acceptance with the probability of Eq. (11) still makes the overall procedure exact. Indeed, in this way the hybrid algorithm of Ref. 5 becomes exact, just as the procedure with a single step makes the Langevin evolution exact. After each accept-reject step, the momentum p are refreshed, their values being replaced by new Gaussian random numbers. The fields ϕ could also be refreshed at this time, or less often, as turns out to be appropriate. The hope is that the microcanonical evolution will sufficiently restrict changes in the action that the final acceptance will remain high for reasonable step sizes.

This procedure contains several parameters which can be adjusted for optimization. First is N_{mic}, the number of microcanonical iterations taken before the global accept-reject step followed by a refreshing of the momenta p. Then there is the step size δ, which presumably should be set to give a reasonable acceptance. Finally, one can also vary the frequency with which the auxiliary scalar fields ϕ are updated.

The arguments for following a microcanonical trajectory for some distance before refreshing the momenta have been stressed in Ref. 5. Reference 16 shows that this approach gives an algorithm where the computer time grows as $V^{5/4}$. For completeness I give another argument for this behavior.

The essence of the approach is to replace a random walk of the A field with a continued motion in the direction of p. As long as the total microcanonical time for a trajectory is smaller than some characteristic time for the system, the net change in A will grow linearly with both N_{mic} and δ; thus, Eq. (35) is replaced by

$$\Delta A \sim N_{mic}\delta \ , \tag{38}$$

which should be valid as long as

$$N_{mic}\delta \lesssim 1 \ . \tag{39}$$

With large N_{mic}, the change in the classical energy will also grow. In any given microcanonical step the energy changes by an amount of order δ^3. For N_{mic} of order δ^{-1}, the total energy change will then be of order δ^2. Because the evolution preserves areas in phase space, Eq. (19) still applies to the overall evolution and I have for the expected energy change

$$\langle H'-H\rangle = \tfrac{1}{2}\langle (H'-H)^2\rangle + O((H'-H)^3) = O(\delta^4) \ . \tag{40}$$

Now if I update V-independent variables together, these positive contributions will coherently add and earlier ar-

guments now give an overall acceptance falling as

$$P_{acc} \sim e^{-CV\delta^4} \ . \tag{41}$$

This means that δ should be taken to decrease with volume as $V^{-1/4}$. Correspondingly, N_{mic} grows as $V^{1/4}$, the maximum allowed by Eq. (39). The final result is that the total time required to obtain a substantially changed lattice grows as

$$T \sim V^{5/4} \ . \tag{42}$$

This may be only an asymptotic statement, valid for systems much larger than the correlation length. The main uncertainty lies in the unknown characteristic time scales that determine the ~ 1 right-hand side of Eq. (39). Nevertheless, the above growth is sufficiently slow that it compels further testing. In the remainder of this paper I discuss the application of these ideas to a simple many-electron system.

IV. A MANY-ELECTRON MODEL

I test the above exact algorithm on the two-dimensional Hubbard model.[17] I work on an N by N square lattice, on each site of which can be electrons of spin up or down. An electron of spin $\sigma \in \{\uparrow, \downarrow\}$ is created or destroyed on site i by the operator $a_{i\sigma}^\dagger$ or $a_{i\sigma}$. These satisfy the canonical anticommutation relations

$$[a_{i\sigma}, a_{j\mu}^\dagger]_+ = \delta_{ij}\delta_{\sigma\mu} \ . \tag{43}$$

The quantum Hamiltonian for the system is

$$\mathcal{H} = -K \sum_{\{ij\},\sigma} a_{i\sigma}^\dagger a_{j\sigma} - \frac{U}{2} \sum_i (a_{i\uparrow}^\dagger a_{i\uparrow} - a_{i\downarrow}^\dagger a_{i\downarrow})^2 + h \sum_i (a_{i\uparrow}^\dagger a_{i\uparrow} - a_{i\downarrow}^\dagger a_{i\downarrow}) + \mu \sum_{i,\sigma} a_{i\sigma}^\dagger a_{i\sigma} \ . \tag{44}$$

Here $\{ij\}$ denotes the set of all nearest-neighbor ordered pairs of sites. (Each unordered pair appears twice in the sum, once for each order.) I refer to the variable K as the hopping parameter, U as the interaction or four-fermion coupling, h as the magnetic field, and μ as the chemical potential. I chose the form of the interaction term to be invariant under a particle-hole transformation. The sign of this term is chosen to represent an effective repulsion when both spin up and spin down occupy the same site.

I am interested in the partition function for this Hamiltonian

$$Z = \mathrm{Tr}\, e^{-\beta\mathcal{H}} \ . \tag{45}$$

To proceed I divide β into N_t "time" slices and write

$$Z = \mathrm{Tr}[(e^{-\beta\mathcal{H}/N_t})^{N_t}] \ . \tag{46}$$

I will manipulate the individual slices separately and make approximations which become exact as N_t goes to infinity. The first such approximation is to separate the hopping term from the others and write

$$e^{-\beta\mathcal{H}/N_t} = \exp\left[\frac{K\beta}{N_t}\sum_{\{ij\},\sigma} a_{i\sigma}^\dagger a_{j\sigma}\right]\exp\left[\frac{U\beta}{2N_t}\sum_i (a_{i\uparrow}^\dagger a_{i\uparrow} - a_{i\downarrow}^\dagger a_{i\downarrow})^2 - \frac{h\beta}{N_t}\sum_i (a_{i\uparrow}^\dagger a_{i\uparrow} - a_{i\downarrow}^\dagger a_{i\downarrow}) - \frac{\mu\beta}{N_t}\sum_{i,\sigma} a_{i\sigma}^\dagger a_{i\sigma}\right] + O(K/N_t^2) \ . \tag{47}$$

All the terms in the final exponent in this equation commute with each other; thus I can perform ordinary arithmetic operations on them. In particular, I can introduce an integration over a set of auxiliary fields $\{A_i\}$ located on the sites so as to eliminate the four-fermion terms from the exponent:

$$e^{-\beta\mathcal{H}/N_t}=(2\pi)^{-N^2/2}\int\{dA\}e^{-\sum A^2/2}\exp\left[\frac{K\beta}{N_t}\sum_{\{ij\},\sigma}a_{i\sigma}^\dagger a_{j\sigma}\right]$$

$$\times\exp\left[\sum_i[(U\beta/N_t)^{1/2}A_i-h\beta/N_t](a_{i\uparrow}^\dagger a_{i\uparrow}-a_{i\downarrow}^\dagger a_{i\downarrow})-\frac{\mu\beta}{N_t}\sum_{i,\sigma}a_{i\sigma}^\dagger a_{i\sigma}\right]+O(K/N_t^2)\ . \tag{48}$$

Here $\{dA\}$ denotes integration over all A fields on the respective time slice.

At this point it is convenient to make a particle-hole transformation on the spin-down electrons. The main reason for this is to clarify the positivity of the final fermion determinant. Thus, I define the creation operator for a spin-up hole as the destruction operator for a spin-down electron times a phase:

$$b_{i\uparrow}^\dagger=(-1)^i a_{i\downarrow}\ . \tag{49}$$

Here $(-1)^i$ means -1 raised to the parity of the site. Using this to eliminate all spin-down operators from Eq. (48) gives

$$e^{-\beta\mathcal{H}/N_t}=(2\pi)^{-N^2/2}\int\{dA\}e^{-\sum A^2/2}\exp\left[\frac{K\beta}{N_t}\sum_{\{ij\}}(a_i^\dagger a_j+b_i^\dagger b_j)\right]$$

$$\times\exp\left[\sum_i[(U\beta/N_t)^{1/2}A_i-h\beta/N_t](a_i^\dagger a_i+b_i^\dagger b_i-1)-\frac{\mu\beta}{N_t}\sum_i(a_i^\dagger a_i-b_i^\dagger b_j+1)\right]+O(K/N_t^2)\ . $$

$$\tag{50}$$

Here I suppress the spin index because they are all up. Note that the integrand factors into pieces involving a and b operators separately. Furthermore, these factors are of the same form if μ vanishes.

Putting the time slices together, shifting the A integrals, and writing things as a product of normal-ordered factors gives

$$Z=\exp[N^2\beta(U/2+h-\mu)](2\pi)^{-N^2N_t/2}$$

$$\times\int(dA)e^{-A^2/2}\mathrm{Tr}\prod_t:\left[\prod_{\{ij\}}[1+(K\beta/N_t)a_i^\dagger a_j]\prod_i(1+a_i^\dagger a_i\{\exp[(U\beta/N_t)^{1/2}A_{i,t}-(h+U+\mu)\beta/N_t]-1\})\right.$$

$$\left.\times(\{a,a^\dagger,\mu\}\rightarrow\{b,b^\dagger,-\mu\})[1+O(K/N_t^2)]\right]:\ , \tag{51}$$

where $(dA)e^{-A^2/2}$ is shorthand for $\prod_{i,t}(dA_{i,t}e^{-A_{i,t}^2/2})$.

The trace of this product of quadratic forms of the fermionic fields can now be evaluated by standard methods.[18-20] For completeness, I present in the Appendix yet another derivation. The final result is

$$Z=\lim_{K/N_t\to0}\exp[N^2\beta(U/2+h-\mu)](2\pi)^{-N^2N_t/2}\int(dA)e^{-A^2/2}\det(M_+)\det(M_-)\ . \tag{52}$$

The matrices M_+ and M_- are easily specified by their matrix elements between arbitrary vectors ψ^* and ψ:

$$\psi^*M_\pm\psi=\frac{K\beta}{N_t}\sum_{\{ij\},t}\psi_{i,t}^*\psi_{j,t}+\sum_{i,t}\psi_{i,t}^*(\psi_{i,t}-\psi_{i,t-1})+\sum_{i,t}\psi_{i,t}^*\psi_{i,t}\{\exp[(U\beta/N_t)^{1/2}A_{i,t}-(h+U\pm\mu)\beta/N_t]-1\}\ . \tag{53}$$

Here antiperiodic boundary conditions are to be taken in the t direction; so, I define $\psi_{i,0}=-\psi_{i,t}$. Note that the finite-N_t approximations have been made in such a manner that the result remains exact when the hopping parameter K vanishes. This is convenient for testing while adding only insignificantly to the required computer time. Note also that the above matrices are not Hermitian.[20]

Substituting $\det M_+\det M_-=\det(M_+M_-^\dagger)$ puts the problem into the form of Eq. (2) where the fermionic matrix is the product $M_+M_-^\dagger$. If this product is a positive

matrix, the entire discussion of Secs. II and III applies. While, in general, positivity is not guaranteed, it is when μ vanishes. In this case, referred to as a half-filled band, M_+ and M_- are identical and $\det(M_+M_-^\dagger)$ is a square. From now on I restrict my discussion to this case with $\mu=0$ and refer to $M_+=M_-$ as simply M. Finding methods to relax the constraint to vanishing chemical potential is an important problem but beyond the scope of this paper.

Applying the discussion of Sec. III, I introduce an auxiliary field ϕ and a momentum field p conjugate to A. I

wish to find configurations of these fields with probability distribution

$$P_{eq}(p, A, \phi) \propto e^{-H(p, A, \phi)} , \tag{54}$$

where

$$H = p^2/2 + A^2/2 + [(MM^\dagger)^{-1}\phi]^2/2 . \tag{55}$$

Here sums over the degrees of freedom are understood.

For the computations discussed below, I took a 6×6 spatial lattice and used $N_t = 8$ time slices. For the updating I used three types of steps. First, to refresh the momenta I merely replace all p with new Gaussian random numbers. Second, to refresh the field ϕ I adapt the prescription of Eqs. (6) and (7) to set

$$\phi = MM^\dagger \chi , \tag{56}$$

where χ represents another Gaussian random vector. Although the p and ϕ updatings could be done at independent times, in the following I refreshed them each simultaneously.

The third part and core of the algorithm, which occurs between these refreshings of p and ϕ, consists of doing N_{mic} iterations of Eq. (37) followed by the global acceptance of the new p and A fields with the probability

$$P_{acc} = \min\left[1, \frac{e^{-H(p', A', \phi)}}{e^{-H(p, A, \phi)}}\right] . \tag{57}$$

This exactly restores detailed balance. For the driving force I take the Langevin choice

$$F_i = -\frac{1}{2}\frac{\partial H}{\partial A_i}$$

$$= -\frac{1}{2}\left[A_i - \left[(MM^\dagger)^{-2}\phi, \frac{\partial M}{\partial A}M^{-1}\phi\right]\right.$$

$$\left. - \left[M^{-1}(MM^\dagger)^{-1}\phi, \frac{\partial M}{\partial A}(MM^\dagger)^{-1}\phi\right]\right] . \tag{58}$$

The appearance of $(MM^\dagger)^{-2}\phi$ in this equation indicates that two conjugate-gradient inversions are required for the calculation of this force. This is true except immediately after refreshing ϕ, in which case $\chi = (MM^\dagger)^{-1}\phi$ is already known. Thus, the algorithm requires $2N_{mic} + 1$ inversions per microcanonical trajectory.

As observables to monitor during the simulations, I measure both the electron density and the pair density

$$\langle n_{i\uparrow} \rangle \equiv \langle a_{i\uparrow}^\dagger a_{i\uparrow} \rangle, \quad \langle n_{i\uparrow}n_{i\downarrow} \rangle \equiv \langle a_{i\uparrow}^\dagger a_{i\uparrow}a_{i\downarrow}^\dagger a_{i\downarrow} \rangle . \tag{59}$$

The observable $\langle n_{i\uparrow}n_{i\downarrow} \rangle$ has true dynamical significance. When the coupling and magnetic field vanish, it should have the value $\frac{1}{4}$, and deviations therefrom are a consequence of the interaction.

Since the band is half-filled, the exact solution with no magnetic field should give $\langle n_{i\uparrow} \rangle = \frac{1}{2}$. However, the finite-$N_t$ approximations break the symmetry between spin up and down and give a slightly smaller number (although I still have $\langle n_{i\uparrow} \rangle = 1 - \langle n_{i\downarrow} \rangle$). Thus, this quanti-

ty is a useful monitor of the finite-N_t errors.

To calculate these observables I relate them to expectation values of matrix elements of M^{-1}. This gives

$$\langle n_{i\uparrow} \rangle = \langle 1 + M_{i,t;i,t+1}^{-1} \rangle ,$$

$$\langle n_{i\uparrow}n_{i\downarrow} \rangle = -\langle M_{i,t;i,t+1}^{-1}(1 + M_{i,t;i,t+1}^{-1}) \rangle . \tag{60}$$

As these formulas are probably not very transparent, I derive them in the Appendix. In the following calculations I average these quantities over all spatial sites on one time slice. Even though this involves N^2 inversions per lattice, most of the computer time is still spent on updating and not on measuring.

For the purposes of testing the algorithm, I set the inverse temperature β and the hopping parameter K to unity. I set the magnetic field h to zero. Of course, I also set the chemical potential μ to zero for the reasons discussed above. As mentioned earlier, I work on a 6×6 lattice with eight time slices.

In Fig. 1 I show the resulting measurements of $\langle n_{i\uparrow} \rangle$ and $\langle n_{i\uparrow}n_{i\downarrow} \rangle$ as a function of the coupling parameter U. Each point was obtained from a run of 128 trajectories of $N_{mic} = 64$ microcanonical steps. The parameter δ was chosen to be $\frac{1}{64}$. Each trajectory was followed by an accept-reject step and then both p and ϕ were refreshed. The observables were measured each 16 trajectories. The starting lattice in each case came from a run of 32 similar trajectories applied to a lattice with A initially zero. Thus, each point in this figure comes from a run of 10 240 microcanonical steps.

As U was increased, the amount of correlation between successive lattices in the Markov chain of updatings increased dramatically. For U larger than 2.4 the above runs were too short for the successive measurements to be considered as independent.

Figure 1 also shows curves representing the exact solution to the theory when the hopping parameter is set to zero. Note how the delocalization allowed by the hop-
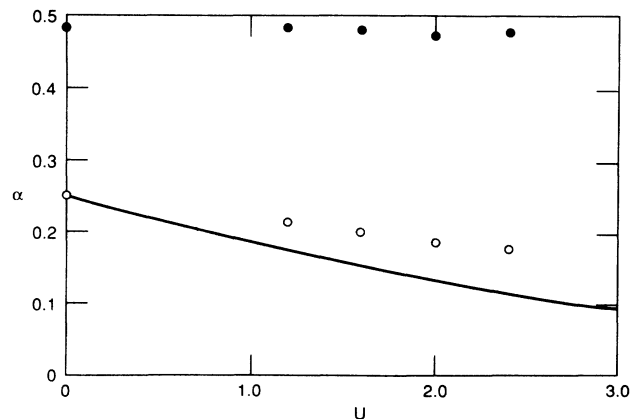

FIG. 1. The behavior of the densities $\langle n_{i\uparrow} \rangle$ (solid points) and $\langle n_{i\uparrow}n_{i\downarrow} \rangle$ (open points) as a function of the coupling U. The other parameters are $\beta = 1$, $K = 1$, and $h = \mu = 0$. The solid curve represents the exact solution for vanishing K. The statistical errors are approximately the size of the points. The fact that $\langle n_{i\uparrow} \rangle$ is less than 0.5 is a consequence of N_t being finite.

ping increases the value of the pair density $\langle n_{i\uparrow}n_{i\downarrow}\rangle$. Note also that the value of $\langle n_{i\uparrow}\rangle$ is measurably less than the value 0.5 expected for infinite N_t. At vanishing U the value expected for $\langle n_{i\uparrow}\rangle$ on this size is 0.485.

Selecting $U=2$ for further analysis, I performed several experiments using as an initial lattice the final lattice from the above run. The purpose of these experiments was to study the parameters $N_{\rm mic}$ and δ. Of course, any conclusions for optimum values of these parameters should depend, perhaps strongly, on the couplings and the value of N_t.

I show in Fig. 2 the effects on the algorithm of varying $N_{\rm mic}$ while holding the product $N_{\rm mic}\delta$ fixed to unity. Thus, the total "physical" time spent on each trajectory is held fixed while the discretization is varied. As a measure of the decorrelation rate, I consider the correlation between the A fields on lattices separated by four trajectories. The correlation between lattices A and A' is defined as

$$C(A,A')=\frac{(A-\langle A\rangle, A'-\langle A'\rangle)}{[(A-\langle A\rangle)^2(A'-\langle A'\rangle)^2]^{1/2}}, \qquad (61)$$
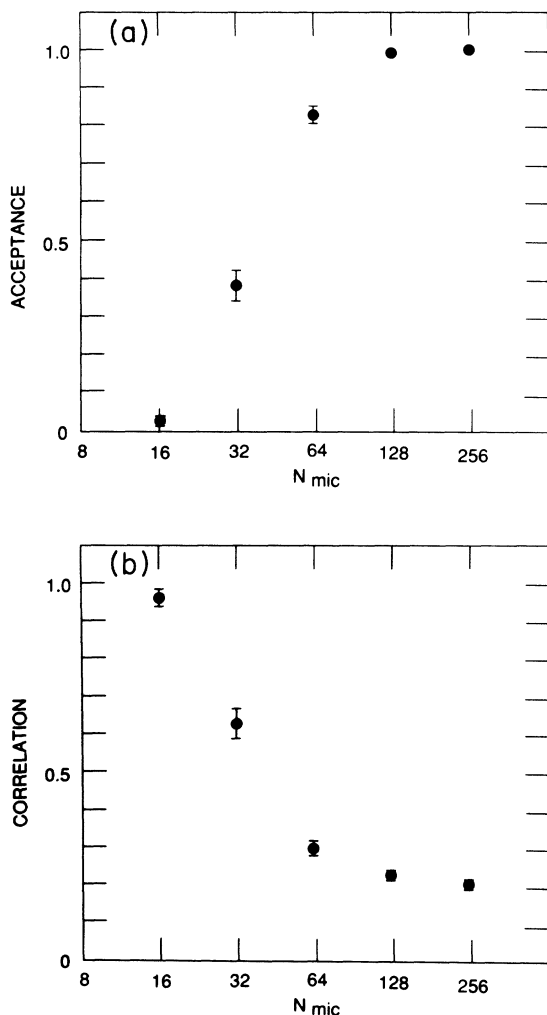
where the inner product is

$$(A,A')=\sum_{i,t}A_{i,t}A'_{i,t}. \qquad (62)$$

As $N_{\rm mic}$ increases, the trajectory becomes more microcanonical and the overall acceptance increases to unity. Meanwhile, the correlation after four trajectories drops towards the value it would have for a true hybrid microcanonical algorithm.[5] I note in passing that when this experiment is done with 32-bit precision, the acceptance remains well below unity, even for quite large $N_{\rm mic}$. Round-off errors at low precision tend to increase the classical energy and suppress acceptance. For this reason all runs presented in this paper use 64-bit arithmetic.

As trajectories with increasing $N_{\rm mic}$ require increasing computer time, one cannot conclude an optimum $N_{\rm mic}$ directly from this figure. In Fig. 3 I show the correlation between lattices separated by $256/N_{\rm mic}$ trajectories while still holding $N_{\rm mic}\delta$ fixed. Thus, these lattices are separated by a constant 256 microcanonical steps and therefore, an approximately constant amount of computer time. Note the minimum at $N_{\rm mic}\sim64$.

In Fig. 4 I consider varying the product $N_{\rm mic}\delta$ while holding $\delta=\frac{1}{64}$ fixed. This corresponds to varying the physical trajectory length. Again I plot the correlation between lattice separated by a fixed number of 256 microcanonical steps. This shows that the choice of $N_{\rm mic}\delta=1$ was not unreasonable, but with hindsight $N_{\rm mic}\delta=2$ would have been better. Indeed, this illustrates a practical difficulty with having too many parameters to optimize. The minimum seen in this figure is presumably partly due to decreasing acceptance with increasing trajectory length and partly due to recurrence in the Hamiltonian dynamics.

V. CONCLUDING REMARKS

I have discussed in some detail exact algorithms for simulating fermionic fields by adding a global accept-reject step to the Langevin and hybrid algorithms. These
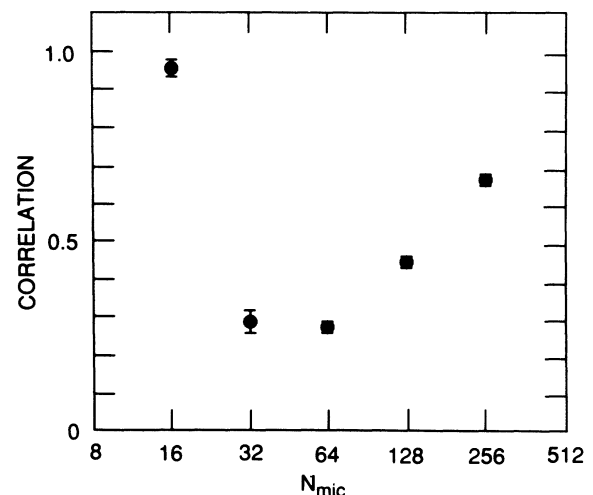

FIG. 2. (a) The global acceptance for trajectories of length $N_{\rm mic}\delta=1$ as a function of $N_{\rm mic}$. (b) The correlation between lattices separated by four of the above trajectories.

FIG. 3. The correlation between lattices separated by 256 microcanonical steps as a function of the parameter $N_{\rm mic}$. The product $N_{\rm mic}\delta$ is held fixed to unity.
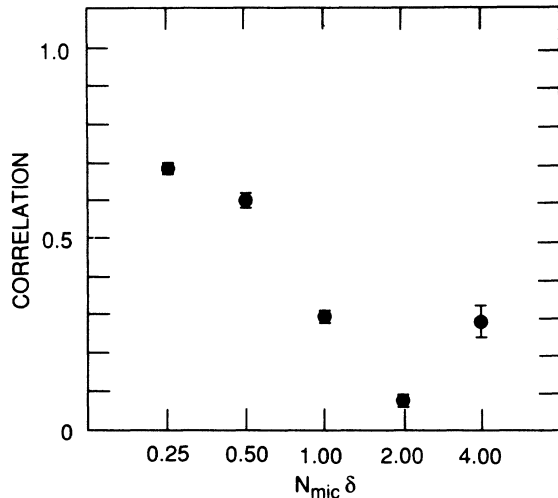
FIG. 4. The correlation between lattices separated by 256 microcanonical steps as a function of the trajectory length $N_{mic}\delta$. Here δ is held fixed to $\frac{1}{64}$.

approaches offer the promise of a rather favorable volume growth of computer time when compared with other exact algorithms.

Several variations of these algorithms are possible. As the global step makes the method exact, one is free to modify the driving force arbitrarily to increase this final acceptance. While the higher-order form suggested in Eq. (33) may be prohibitively difficult to evaluate, simpler empirical forms could prove useful. One potential time-saving possibility is to only perform a minimal number of iterations of the conjugate-gradient algorithm when calculating the force for the intermediate updates, and only do an accurate evaluation of the action for the final global accept stage. The fact that rather large values for N_{mic} are called for make these savings particularly attractive.

Tests of this idea at $U=2$ indicate no significant loss in decorrelation rate when the number of conjugate-gradient iterations at intermediate stages is restricted to a maximum of 64. For comparison, my stopping criterion for the earlier runs was the accumulation of at least three iterations which do not decrease the quadratic form being explicitly minimized by the conjugate-gradient algorithm. For $U=2$ this required of order 160 iterations with an unbiased initial guess for the inverse. (Using the previous inverse for a starting vector reduces the latter number by a few percent, and extrapolating from the previous two inverses reduces it by a further few percent. Maintaining reversibility with a restricted number of iterations requires unbiased starting conditions.) Thus, I find that restricting the accuracy of the inversions for intermediate steps gains more than a factor of 2 in computer time when compared to doing all inversions with the above stopping criterion. Of course, all these numbers depend on the couplings.

The Hubbard model tests indicate the importance of taking a rather small value for the step size δ. This is somewhat surprising because Eq. (32) indicates that the
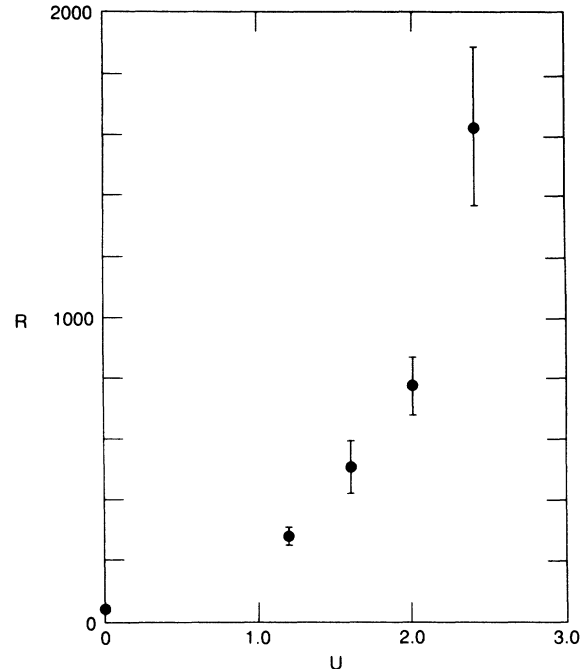

FIG. 5. The ratio of the largest to the smallest eigenvalue of $MM^\dagger$ as a function of the coupling U.

coherent suppression of acceptance does not occur until order δ^4. For $\delta = \frac{1}{64}$ this is a particularly small number; so, there must be some large compensating number in the physics of the system. To try to find a quantity, I investigated the largest and smallest eigenvalues of the matrix $MM^\dagger$ on the lattices used for the points in Fig. 1. To find these eigenvalues I applied $MM^\dagger$ or $(MM^\dagger)^{-1}$ 20 times to a random vector. While this procedure in principle gives only a bound on the eigenvalues, increasing the number of applications of the corresponding matrix did not significantly change the result. I show in Fig. 5 the average ratio of these eigenvalues as a function of U. The fact that this ratio is of order 10^3 shows that the matrix $MM^\dagger$ is rather poorly conditioned and may be the source of the necessity for small δ. Note that this eigenvalue ratio is rapidly rising as U increases. The growth is due to a simultaneous increase of the largest eigenvalue and decrease of the smallest. This rise is presumably the reason that convergence of the algorithm deteriorated rapidly as U became larger than 2.4. In addition to lattices becoming decorrelated more slowly at larger U, the number of conjugate-gradient steps required for the matrix inversion also increases.

ACKNOWLEDGMENT

APPENDIX

In this appendix I study traces of products of normal-ordered fermionic operators and show how to convert

them to integrals over anticommuting variables. This allows me to derive Eq. (52). Inserting the corresponding observables into the normal-ordered product gives Eq. (60).

To begin I consider the natural mapping between the Hilbert space of fermionic states and the space of functions of a set of anticommuting variables $\{\psi^*\}$. Consider an arbitrary state $|f\rangle$ obtained by applying a function of creation operators to the vacuum state

$$|f\rangle = f(a^\dagger)|0\rangle . \tag{A1}$$

The one-to-one-mapping of this state onto functions of the ψ^* is simply

$$|f\rangle \leftrightarrow f(\psi^*) . \tag{A2}$$

With this correspondence, I clearly have

$$a_i^\dagger|f\rangle \leftrightarrow \psi_i^* f(\psi^*) . \tag{A3}$$

Formally defining differentiation with respect to anticommuting numbers, I write

$$a_i|f\rangle \leftrightarrow \frac{d}{d\psi^*}f(\psi^*) . \tag{A4}$$

To put this in a more convenient form, I introduce a Dirac δ function for anticommuting numbers

$$\delta(\psi^*,\psi'^*) = \int (d\psi)e^{\sum(\psi^*-\psi'^*)\psi} . \tag{A5}$$

Here $(d\psi)$ denotes $\prod d\psi_i$ in some prescribed order, and integration is defined as usual to be a linear mapping with

my phase conventions being

$$\int d\psi_j\psi_j = i, \quad \int d\psi_j 1 = 0 . \tag{A6}$$

The Dirac δ function has the appropriate property

$$f(\psi^*) = \int \delta(\psi^*,\psi'^*)(d\psi')^* f(\psi'^*) . \tag{A7}$$

Combining the above equations, we have the correspondences

$$\begin{aligned} a_i|f\rangle &\leftrightarrow \int \psi_i(d\psi)e^{\sum(\psi^*-\psi'^*)\psi}(d\psi')^* f(\psi'^*) , \\ a_i^\dagger|f\rangle &\leftrightarrow \int \psi_i^*(d\psi)e^{\sum(\psi^*-\psi'^*)\psi}(d\psi')^* f(\psi'^*) . \end{aligned} \tag{A8}$$

Thus, we have the simple result that a_i and $a_i^\dagger$ correspond, respectively, to ψ_i and ψ_i^* under the above integral. This immediately generalizes to any normal-ordered function of a_i and $a_i^\dagger$:

$$:g(a^\dagger,a):|f\rangle$$
$$\leftrightarrow \int g(\psi^*,\psi)(d\psi)e^{\sum(\psi^*-\psi'^*)\psi}(d\psi')^* f(\psi'^*) . \tag{A9}$$

It is important that the function g be normal ordered. This is because of the correspondence of a with a derivative as exhibited in Eq. (A4). If factors of a lie to the left of factors of $a^\dagger$, this derivative will pick up additional contributions. Remember, ψ and ψ^* simply anticommute while a and $a^\dagger$ have a Kronecker δ in their anticommutation relation.

If I consider a product of normal-ordered factors, Eq. (A9) iterates to give

$$:g_1(a^\dagger,a)::g_2(a^\dagger,a): \cdots :g_{N_t}(a^\dagger,a): |f\rangle$$

$$\leftrightarrow \int g_1(\psi^*,\psi_1)(d\psi_1)e^{\sum(\psi^*-\psi_2^*)\psi_1}(d\psi_2)^* g_2(\psi_2^*,\psi_2)(d\psi_2)e^{\sum(\psi_2^*-\psi_3^*)\psi_2}(d\psi_3)^* \times \cdots$$

$$\times g_{N_t}(\psi_{N_t}^*,\psi_{N_t})(d\psi_{N_t})e^{\sum(\psi_{N_t}^*-\psi'^*)\psi_{N_t}}(d\psi')^* f(\psi'^*) . \tag{A10}$$

The trace of a normal-ordered operator can be written in the form of an integral over anticommuting numbers

$$\mathrm{Tr}:g(a^\dagger,a): = \int (d\psi)^* g(\psi^*,\psi)e^{2\sum\psi^*\psi}(d\psi) . \tag{A11}$$

For each mode, expanding the exponential gives two terms representing whether or not that mode is occupied. The factors in this equation are easily verified on a small system.

I can now combine Eqs. (A9)–(A11) to reexpress the trace of a product of normal-ordered operators as a multiple integral over anticommuting numbers. A little algebra gives

$$\mathrm{Tr}[:g_1(a^\dagger,a)::g_2(a^\dagger,a): \cdots :g_{N_t}(a^\dagger,a):] = \int \prod_{t=1}^{n} [(d\psi_t)^* g_t(\psi_t^*,\psi_t)(d\psi_t)e^{\sum\psi_t^*(\psi_t-\psi_{t-1})}] \Big|_{\psi_0=-\psi_{N_t}} . \tag{A12}$$

Note how the exponential in Eq. (A11) serves to flip a sign and give the well-known antiperiodic boundary condition.

Formula (A12) is exactly what is needed to convert Eq. (51) into a product of integrals of a form which can be done using the standard Matthews-Salam[14] formula

$$\int (d\psi)^*(d\psi)e^{\psi^* M\psi} = \det M . \tag{A13}$$

This directly gives Eqs. (52) with the matrices given in Eq. (53).

Now I turn to the expectation of a normal-ordered observable $:h(a^\dagger,a):$. Such an expectation value just inserts another factor into the above trace of a product of operators. This gives another integration, inserting into Eq. (A12) the following additional factors:

$$\int (d\psi_{N_t+1})^*(d\psi_{N_t+1})h(\psi_{N_t+1}^*,\psi_{N_t+1})$$

$$\times \exp[\psi_{N_t+1}^*(\psi_{N_t+1}-\psi_{N_t})+\psi_1^*(\psi_{N_t+1}-\psi_{N_t})] . \tag{A14}$$

Specializing to the case

$$:h(a^\dagger,a):=\langle n \rangle = a^\dagger a \ , \tag{A15}$$

I can explicitly do the extra integrations in Eq. (A14), reducing the insertion to

$$1-\psi_1^*\psi_{N_t} \ . \tag{A16}$$

Putting this into the Gaussian integral of Eq. (A13) and cylicly shifting one unit in the time direction gives the first half of Eq. (60). The second half of that equation follows from similar arguments after doing the particle-hole transformation on the spin-down components.

The temporal splitting in the insertion of Eq. (A16) is somewhat nonintuitive. A first guess might have been to merely insert $\psi_1^*\psi_1$. This is almost correct, but differs from the above by terms of order $1/N_t$. Indeed, these objects are related by what can be regarded as an equation of motion

$$\langle 1-\psi_1^*\psi_{N_t} \rangle = \left\langle \psi_1^*\psi_1 - \frac{\beta}{N_t}\psi_1^*\frac{d}{d\psi_1^*}\mathcal{H}(\psi_1^*,\psi_1)\right\rangle \ . \tag{A17}$$

Thus, if we ignore $1/N_t$ errors, either form is acceptable. Caution may be necessary when considering field theories where the extra term may have divergence giving rise to anomalies surviving in the continuous time limit.

[1]F. Fucito, E. Marinari, G. Parisi, and C. Rebbi, Nucl. Phys. B180 [FS2], 369 (1981).

[2]J. Polonyi and H. W. Wyld, Phys. Rev. Lett. 51, 2257 (1983); J. Kogut, J. Polonyi, H. W. Wyld, and D. K. Sinclair, ibid. 54, 1475 (1985).

[3]G. G. Batrouni, G. R. Katz, A. S. Kronfeld, G. P. Lepage, B. Svetitsky, and K. G. Wilson, Phys. Rev. D 32, 2736 (1985).

[4]A. Ukawa and M. Fukugita, Phys. Rev. Lett. 55, 1854 (1985).

[5]S. Duane and J. Kogut, Phys. Rev. Lett. 55, 2774 (1985); Nucl. Phys. B275, 398 (1986).

[6]D. Weingarten and D. Petcher, Phys. Lett. 99B, 333 (1981).

[7]R. T. Scalettar, D. J. Scalapino, and R. L. Sugar, Phys. Rev. B 34, 7911 (1986).

[8]S. Gottlieb, W. Liu, D. Toussaint, and R. L. Sugar, Phys. Rev. D 35, 2611 (1986).

[9]S. Duane, A. D. Kennedy, B. J. Pendleton, and D. Roweth, Phys. Lett. B 195, 2 (1987).

[10]H. Gausterer and S. Sanielevici, BNL report, 1988 (unpublished).

[11]R. Gupta, in Field Theory of the Lattice, proceedings of the International Symposium, Seillac, France, 1987, edited by A. Billoire et al. (Nucl. Phys. B, Proc. Suppl. Vol. 4) (North-Holland, Amsterdam, 1988).

[12]K. Bitar, R. Horsley, A. Kennedy, S. Meyer, and P. Rossi, Report No. FSU-SCRI-88-14, 1988 (unpublished).

[13]N. Metropolis, A. W. Rosenbluth, M. N. Rosenbluth, A. H. Teller, and E. Teller, J. Chem. Phys. 21, 1087 (1953).

[14]P. T. Matthews and A. Salam, Nuovo Cimento 12, 563 (1954).

[15]M. Creutz, in Lattice Gauge Theory Using Parallel Processors, edited by X. Li, Z. Qiu, and H.-C. Ren (Gordon and Breach, New York, 1987), p. 105.

[16]R. Gupta, G. Kilcup, and S. Sharpe, this issue, Phys. Rev. D 38, 1278 (1988). I am particularly grateful to S. Sharpe for pointing out an incorrect argument in the early draft of my paper.

[17]J. Hubbard, Proc. R. Soc. London A276, 283 (1963); A281, 401 (1964).

[18]R. Blankenbecler, D. Scalapino, and R. Sugar, Phys. Rev. D 24, 4295 (1981).

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

[20]M. Creutz, Phys. Rev. D 35, 1460 (1987).