Monte Carlo Algorithm

In this section we describe the algorithms that we use to perform Quantum Monte Carlo simulations of systems with fermion degrees of freedom. We have found it most useful to carry out simulations within the grand-canonical ensemble. Zero-temperature results can be obtained from such simulations by extrapolating low-temperature results. The average value of a physical observable, O, in the grand-canonical ensemble is given by

$$\langle O \rangle = \frac{\mathrm{Tr}\left(O\, e^{-\beta H}\right)}{\mathrm{Tr}\left(e^{-\beta H}\right)}, \tag{1}$$

where β is the inverse temperature and H is the Hamiltonian. Our objective is to evaluate $\langle O \rangle$ numerically.

Let us begin by briefly reviewing the standard approach to simulating systems with boson degrees of freedom. We will need the results for our discussion of electron-phonon systems, and it will be instructive to contrast the boson and fermion algorithms. for concreteness we take the Hamiltonian to have the form

$$H = \tfrac{1}{2}\sum_{i=1}^{N} p_i^2 + V(x_1, \ldots, x_N) = K + V. \tag{2}$$

Here N is the number of spatial lattice points. The first step is to express the traces in Eq. (1) as path integrals. To this end we divide the line 0 to β into L segments of width $\Delta\tau$, so that $\beta = L\Delta\tau$. We shall refer to these segments as time slices, and define the Euclidean-time, τ_ℓ, to be $\tau_\ell = \ell\Delta\tau$. Making use of the fact that

$$e^{-\Delta\tau H} = e^{-\frac{1}{2}\Delta\tau K}\, e^{-\Delta\tau V}\, e^{-\frac{1}{2}\Delta\tau K} + O\left(\Delta\tau^3\right), \tag{3}$$

we can use the Trotter formula to write the partition function in the form

$$Z = \mathrm{Tr}\left(e^{-\beta H}\right) \simeq \int_{-\infty}^{\infty} dp_1 \, \langle p_1 | \, \left[e^{-\frac{1}{2}\Delta\tau K}\, e^{-\Delta\tau V}\, e^{-\frac{1}{2}\Delta\tau K}\right]^l \, |p_1\rangle. \tag{4}$$

The $|p_1\rangle$ are eigenstates of the momentum operators. Since errors of order $\Delta\tau^3$ are made at each of L time slices, the total error introduced into the partition function and $\langle O \rangle$ is of order

$\Delta\tau^2$. This is the only source of systematic error in the calculation. Introducing a complete set of eigenstates of the momentum and coordinate operator at each time slice reduces Eq. (4) to a multidimensional integral. The integration over the momentum variables can be performed at once, yielding

$$Z = \int \delta x \, e^{-S(x)}, \tag{5}$$

where the Euclidean action $S(x)$ is given by

$$S(x) = \Delta\tau \sum_{\ell=1}^{L} \sum_{i=1}^{N} \left(\frac{x_{i,\ell+1} - x_{i,\ell}}{\Delta\tau} \right)^2 + \Delta\tau \sum_{\ell=1}^{L} V(x_{1,\ell}, \ldots, x_{N,\ell}), \tag{6}$$

and δx indicates an integral over all of the coordinates, $x_{i,\ell}$, $i = 1, \ldots, N$, and $\ell = 1, \ldots, L$. We have absorbed a meaningless constant into the integration measure. A similar expression holds for the trace in the numerator of Eq. (1). One is ordinarily interested in operators that depend on the coordinates, in which case

$$\langle O \rangle = \frac{\int \delta x \, O(x) e^{-S(x)}}{\int \delta x \, e^{-S(x)}}. \tag{7}$$

The ratio of integrals in Eq. (7) can be evaluated by importance sampling. If one generates a sequence of coordinate configurations with a probability distribution

$$P(x) = \frac{e^{-S(x)}}{\int \delta x' \, e^{-S(x')}}, \tag{8}$$

then $\langle O \rangle$ is simply the average value of $O(x)$ in these configurations. To generate the configurations, one sweeps through the lattice repeatedly, changing one of the $x_{i,\ell}$ at a time, according to a standard importance sampling prescription, such as the Metropoulos algorithm. In all of the importance sampling algorithms it is necessary to calculate the change in the Euclidean action, $S(x)$, in response to a change in $x_{i,\ell}$. As long as the potential $V(x_{1,\ell}, \ldots, x_{N,\ell})$ only involves short-range interactions, the number of numerical operations needed to update all of the coordinates once is proportional to NL. Thus, in the absence of critical slowing down, the computer time needed for this approach will grow linearly with the spatial volume and β.

Structure of the Fermion Determinant

Let us now turn to the problem of performing simulations of systems with fermions degrees of freedom. The traces of Eq. (1) can again be expressed as path integrals; however, fermion degrees of freedom are represented in path integrals by anticommuting c-numbers, so these path integrals are not directly susceptible to numerical evaluation. Except for problems in one spatial dimension, it appears to be necessary to perform the trace over the fermion degrees of freedom before numerical methods can be applied. This can be done if the Hamiltonian is quadratic in the fermion creation and annihilation operators, or can be made so by a Hubbard-Stratonovich transformation.

To be concrete, we will use the Hubbard model as an example of our approach. The extension to other models is straightforward. The Hamiltonian for the Hubbard model is

$$\begin{aligned} H = & -t \sum_{\langle ij \rangle, \sigma} \left(c_{i\sigma}^\dagger c_{j\sigma} + c_{i\sigma}^\dagger c_{j\sigma} \right) - \mu \sum_i (n_{i+} + n_{i-}) \\ & + U \sum_i \left(n_{i+} - \tfrac{1}{2} \right) \left(n_{i-} - \tfrac{1}{2} \right). \end{aligned} \tag{9}$$

Here $c_{i\sigma}^\dagger$ and $c_{i\sigma}$ are the creation and annihilation operators for electrons with z component of spin σ at lattice site i, and $n_{i\sigma} = c_{i\sigma}^\dagger c_{i\sigma}$. The sum $\langle i, j \rangle$ is over all pairs of nearest neighbor lattice sites. t is the hopping parameter, μ the chemical potential, and U the Coulomb coupling constant, which is taken to be positive. If we define the kinetic energy operator by

$$\begin{aligned} K &= -t \sum_{\langle i,j \rangle, \sigma} \left(c_{i\sigma}^\dagger c_{j\sigma} + c_{i\sigma}^\dagger c_{j\sigma} \right) - \mu \sum_i (n_{i+} + n_{i-}) \\ &= \sum_{ij\sigma} c_{i\sigma}^\dagger k_{ij} c_{j\sigma}, \end{aligned} \tag{10}$$

and the potential energy operator by $V = H - K$, then making use of Eq. (3), we can write the partition function in the form

$$Z = \mathrm{Tr}\, e^{-\beta H}$$

$$\simeq \mathrm{Tr}\left[e^{-\frac{1}{2}\Delta\tau K}\, e^{-\Delta\tau V}\, e^{-\frac{1}{2}\Delta\tau K}\right]^L = \mathrm{Tr}\left[e^{-\Delta\tau V}\, e^{-\Delta\tau K}\right]. \tag{11}$$

In Eq. (11) we have made use of the invariance of the trace under cyclic permutations of the operators. Again the error in Z introduced by the Trotter approximation is of order $\Delta\tau^2$. We can perform a similar manipulation of the trace in the numerator of Eq. (1),

$$\mathrm{Tr}\left(O\, e^{-\beta H}\right) \simeq \mathrm{Tr}\left(O\left[e^{-\frac{1}{2}\Delta\tau K}\, e^{-\Delta\tau V}\, e^{-\frac{1}{2}\Delta\tau K}\right]^L\right), \tag{12}$$

with errors once more of order $\Delta\tau^2$. We now replace the operator O in the right-hand side of Eq. (12) by $\exp(-\frac{1}{2}\Delta\tau K) O \exp(\frac{1}{2}\Delta\tau K)$. It might appear that this substitution would introduce an error of order $\Delta\tau$ given by

$$\Delta = \tfrac{1}{2}\Delta\tau \mathrm{Tr}\left([O,K]\left[e^{-\frac{1}{2}\Delta\tau K}\, e^{-\Delta\tau V}\, e^{-\frac{1}{2}\Delta\tau K}\right]^L\right). \tag{13}$$

However, Δ vanishes. To see this, we note that in order to carry out a Monte Carlo simulation we require a basis in which the matrix elements of K and V are real. The eigenstates of the operators $n_{i\sigma}$ provide one such basis for all models that we have studied to date. The operators O in which we are interested also have real matrix elements in this basis. So, by judiciously inserting complete sets of these states, the trace in Eq. (13) can be expressed as a sum over real matrix elements. As a result, Δ must be real. On the other hand, since O, K and V are hermitian operators, the trace is of a product of a hermitian and an antihermitian operator, and therefore pure imaginary. Thus, $\Delta = 0$, and we can write

$$\mathrm{Tr}\left(O\, e^{-\beta H}\right) \simeq \mathrm{Tr}\left(O\left[e^{-\Delta\tau V}\, e^{-\Delta\tau K}\right]^L\right), \tag{14}$$

with errors of order $\Delta\tau^2$. We shall also be interested in evaluating correlation functions of the form $\langle O(\tau)^\dagger O(0)\rangle$, where O may or may not be hermitian. Here

$$O(\tau) = e^{\tau H} O\, e^{-\tau H}. \tag{15}$$

A trivial generalization of the argument leading to Eq. (14) shows that

$$\mathrm{Tr}\left(e^{-\beta H}O(\tau)^{\dagger}O(0)\right) \simeq \mathrm{Tr}\left(\left[e^{-\Delta\tau V}\,e^{-\Delta\tau K}\right]^{L_1} O^{\dagger}\,\left[e^{-\Delta\tau V}\,e^{-\Delta\tau K}\right]^{L_2} O\right), \tag{16}$$

with $L_1\Delta\tau = \beta - \tau$ and $L_2\Delta\tau = \tau$. Again the errors are of order $\Delta\tau^2$. The somewhat surprising conclusion is that the Trotter breakup

$$e^{-\Delta\tau H} \simeq e^{-\Delta\tau V}\,e^{-\Delta\tau K}, \tag{17}$$

which itself introduces $\Delta\tau^2$ errors at each time slice, only introduces $\Delta\tau^2$ errors into the partition function and expectation values of interest. We therefore use this breakup in our simulations.

We now turn to the problem of performing the trace over the electron degrees of freedom. This can be done analytically provided the operators K and V are quadratic in the electron creation and annihilation operators. This will always be the case for the kinetic energy operator. for some models, such as those describing electron-phonon interactions it is also true for the potential energy operator. When V is not quadratic in the electron creation and annihilation operators, it can ordinarily be made so through a Hubbard-Stratonovich transformation. The Hubbard model is a case in point. ¿From Eq. (9) we see that for this model V is a sum of commuting single site operators. We can therefore write

$$e^{-\Delta\tau V} = \prod_{i=1}^{N} e^{-\Delta\tau U(n_{i+}-\frac{1}{2})(n_{i-}-\frac{1}{2})}. \tag{18}$$

Thus, by making use of the identity

$$e^{-\Delta\tau U\left(n_{i+}-\frac{1}{2}\right)\left(n_{i-}-\frac{1}{2}\right)} = \left(\frac{\Delta\tau}{\pi}\right)^{\frac{1}{2}} e^{-\frac{1}{4}\Delta\tau U}\int_{-\infty}^{+\infty} dx_{i,\ell}\, e^{-\Delta\tau\left[x_{i,\ell}^2+(2U)^{\frac{1}{2}}x_{i,\ell}(n_{i+}-n_{i-})\right]} \tag{19}$$

at every lattice point i on every time slice ℓ, we transform the quartic interaction between the spin-up and -down electrons into a quadratic interaction in which electrons in each

spin state interact with the field $x_{i,\ell}$, which plays the role of a Euclidean-time dependent magnetic field. The potential energy operator for the $\ell^{\rm th}$ time slice is

$$\mathcal{U}_\ell (2\,U)^{\frac{1}{2}} \sum_i x_{i,\ell} (n_{i+} - n_{i-}). \tag{20}$$

The price that we pay for this simplification is the integrals over the $x_{i,\ell}$.

The continuous Hubbard-Stratonovich transformation generated by Eq. (19) is required for algorithms based on the Langevin equation or molecular dynamics. For standard Monte Carlo simulations one may alternatively use Hirsch's discrete Hubbard-Stratonovich transformation. Hirsch notes that

$$e^{-\Delta\tau U \left(n_{i+} - \frac{1}{2}\right)\left(n_{i-} - \frac{1}{2}\right)} = \tfrac{1}{2}\, e^{-\frac{1}{4}\Delta\tau U} \sum_{s_{i,\ell}=\pm 1} e^{-\Delta\tau s_{i,\ell}\lambda (n_{i+} - n_{i-})}, \tag{21}$$

where $\cosh(\Delta\tau\lambda) = \exp(\Delta\tau U/2)$. Proceeding as in the case of the continuous transformation yields a potential energy operator for the $\ell^{\rm th}$ time slice

$$\mathcal{U}_\ell = \sum_i \lambda s_{i,\ell}\, (n_{i+} - n_{i-}) = \sum_{ij\sigma} c^\dagger_{i\sigma} u^\sigma_{\ell\, i,j} c_{j\sigma}. \tag{22}$$

The discrete Hubbard-Stratonovich transformation has the advantage of simplicity in that one has only to sum over the spin variables $s_{i,\ell} = \pm 1$. In addition, it appears to lead to a faster exploration of configuration space in the simulations.

Having reduced the potential energy operator to a quadratic form in the electron creation and annihilation operators, it is possible to perform the trace over the electron degrees of freedom for an individual configuration of the Hubbard-Stratonovich variables. One can then perform the integrals or sums over these variables by importance sampling techniques. For all models that we have studied to date, the traces factorize into a product of traces for the spin-up and spin-down electrons, and we shall assume that is the case in our discussion.

The traces can be performed straightforwardly using the coherent state representation. We write

$$|\alpha, \sigma\rangle = e^{c^\dagger_\sigma \cdot \alpha} |0\rangle \qquad\qquad \langle\bar{\alpha}, \sigma| = \langle 0| e^{\bar{\alpha}\cdot c_\sigma}, \tag{23}$$

where $|0\rangle$ is the vacuum state, $c_{i\sigma}|0\rangle = 0$, and α and $\bar{\alpha}$ are vectors of anticommuting c-numbers. We have adopted vector notation, so, for example,

$$c_\sigma^\dagger \cdot \alpha = \sum_{i=1}^{N} c_{i\sigma}^\dagger \alpha_i. \tag{24}$$

The states $|\alpha, \sigma\rangle$ and $\langle\bar{\alpha}, \sigma|$ are eigenstates of the electron annihilation and creation operators respectively:

$$\begin{aligned} c_{i\sigma}|\alpha, \sigma\rangle &= \alpha_i |\alpha, \sigma\rangle \\ \langle\bar{\alpha}, \sigma| c_{i,\sigma}^\dagger &= \langle\bar{\alpha}, \sigma|\bar{\alpha}_i. \end{aligned} \tag{25}$$

As a result,

$$\langle\bar{\alpha}, \sigma|\alpha, \sigma\rangle = e^{\bar{\alpha}\cdot\alpha}. \tag{26}$$

Making use of the standard rules of integration for Grassman variables,

$$\int d\alpha_i = 0 \qquad \int d\alpha_i \, \alpha + i = 1, \tag{27}$$

we see that

$$\begin{aligned} \mathcal{Z}_\sigma(s) &= \mathrm{Tr}\left(e^{-\Delta\tau c_\sigma^\dagger \cdot u_L^\sigma \cdot c_\sigma} \, e^{-\Delta\tau c_\sigma^\dagger \cdot k \cdot c_\sigma} \ldots e^{-\Delta\tau c_\sigma^\dagger \cdot u_1^\sigma \cdot c_\sigma} \, e^{-\Delta\tau c_\sigma^\dagger \cdot k \cdot c_\sigma} \right) \\ &= \int \prod_{i=1}^{N} d\alpha_i \, d\bar{\alpha}_i \, e^{\bar{\alpha}\cdot\alpha} \langle\bar{\alpha}, \sigma| e^{-\Delta\tau c_\sigma^\dagger \cdot u_L^\sigma \cdot c_\sigma} \, e^{-\Delta\tau c_\sigma^\dagger \cdot k \cdot c_\sigma} \ldots e^{-\Delta\tau c_\sigma^\dagger \cdot u_1^\sigma \cdot c_\sigma} \, e^{-\Delta\tau_\sigma^\dagger \cdot k \cdot c_\sigma} |\alpha, \sigma\rangle, \end{aligned} \tag{28}$$

where k and u_ℓ^σ are given for the Hubbard model in Eqs. (10, (2), and (22). Using the identity

$$e^{-\Delta\tau c_\sigma^\dagger \cdot h \cdot c_\sigma} c_{i\sigma}^\dagger = \left(c_\sigma^\dagger \cdot e^{-\Delta\tau h} \right)_i e^{-\Delta\tau c_\sigma^\dagger \cdot h \cdot c_\sigma}, \tag{29}$$

where h stands for either k or u_ℓ^σ, we see that

$$\mathcal{Z}_\sigma(s) = \int \prod_{i=1}^{L} d\alpha_i \, d\bar{\alpha}_i \, \langle\bar{\alpha}, \sigma| B_L^\sigma \ldots B_1^\sigma \alpha, \sigma\rangle, \tag{30}$$

with

$$B_\ell^\sigma = e^{-\Delta\tau u_\ell^\sigma} \, e^{-\Delta\tau k}. \tag{31}$$

Then using Eqs. (26) and (27), we obtain

$$\begin{aligned}\mathcal{Z}_\sigma(s) &= \int \prod_{i=1}^{N} d\alpha_i \, d\bar{\alpha}_i \, e^{\bar{\alpha}\cdot[I+B_L^\sigma \ldots B_1^\sigma]\alpha} \\ &= \det[I + B_L^\sigma \ldots B_1^\sigma], \end{aligned} \tag{32}$$

with I denoting the unit matrix. $\mathcal{Z}_\sigma(s)$ is the partition function for an electron with z-component of spin σ propagating through the Hubbard-Stratonovich background field. If the Hubbard-Stratonovich variables did not depend on Euclidian time, then we could undo the breakup of kinetic and potential energy matrices, and find

$$\mathcal{Z}_\sigma(s) = \det\left[I + e^{-\beta h^\sigma}\right] = \prod_{i=1}^{N} \left(1 + e^{-\beta \varepsilon_i^\sigma}\right). \tag{33}$$

Here $h^\sigma = k + u^\sigma$, and ε_i^σ are the eigenvalues of the matrix h^σ. Of course, Eq. (33) is just the standard expression for the partition function of an electron in an external field. Equation (32) is simply the generalization of Eq. (33) to the case of a Euclidean time-dependent field.

Although some quantities of physical interest can be expressed in terms of correlation functions of the Hubbard-Stratonovich variables, we most often want to calculate correlation functions involving the electron creation and annihilation operators. For example, the single-particle, finite-temperature Green's function is given by

$$G_{i,j}^\sigma(\tau_\ell, \tau_{\ell'}) = \langle T[c_{i\sigma}(\tau_\ell) c_{j\sigma}^\dagger(\tau_{\ell'})]\rangle, \tag{34}$$

where T indicates a Euclidean time ordering, and the τ dependence of the creation and annihilation operators is given by Eq. (15). To evaluate such correlation functions it is convenient to make use of the generating functional for a fixed Hubbard-Stratonovich configuration

$$\mathcal{G}^\sigma(\eta, \bar{\eta}, s) = \mathcal{Z}_\sigma(s)^{-1} \mathrm{Tr}\left[e^{-\Delta\tau u_L^\sigma} e^{-\Delta\tau k} e^{\bar{\eta}_{L-1}} \ldots e^{-\Delta\tau u_1^\sigma} e^{-\Delta\tau k} e^{\bar{\eta}_0 \cdot c_\sigma} e^{c_\sigma^\dagger \cdot \eta_0}\right]. \tag{35}$$

η_ℓ and $\bar{\eta}_\ell$ are N-dimensional vectors of anticommuting c-numbers, which serve as sources and sinks for the electrons. The trace in Eq. (35) can easily be evaluated using the coherent

state representation. It is convenient to move the quantities $\exp(c_\sigma^\dagger \cdot \eta_\ell)$ to the far left of the trace, and the quantities $\exp(\bar{\eta} \cdot c_\sigma)$ to the far right, using Eq. (29), so that c_σ and $c_\sigma^\dagger$ can act on the coherent state ket and bra, respectively. We then find

$$\mathcal{G}^\sigma(\eta,\bar{\eta},s) = \mathcal{Z}_\sigma(s)^{-1} \int \prod_{i=1}^{N} d\alpha_i \, d\bar{\alpha}_i \, \exp(\bar{\alpha}[I + B_L^\sigma \ldots B_1^\sigma]\alpha) \, \exp\left(\sum_{\ell=0}^{L-1} \bar{\eta}_\ell B_\ell^\sigma \ldots B_1^\sigma \alpha\right)$$
$$\times \exp\left(\sum_{\ell=0}^{L-1} \bar{\alpha} B_L^\sigma \ldots B_{\ell+1}^\sigma \eta_\ell\right) \, \exp\left(\sum_{\ell \geq \ell'} \bar{\eta}_\ell B_\ell^\sigma \ldots B_{\ell'+1}^\sigma \eta_{\ell'}\right). \tag{36}$$

The term $\bar{\eta}_\ell B_\ell^\sigma \ldots B_1^\sigma \alpha$ reduces to $\bar{\eta}_0 \alpha$ for $\ell = 0$, while $\bar{\eta}_\ell B_\ell^\sigma \ldots B_{\ell'+1}^\sigma \eta_{\ell'}$ reduces to $\bar{\eta}_\ell \eta_\ell$ for $\ell = \ell'$. since integrals over Grassman variables are invariant under a shift by an anticommuting c-number, the integrals in Eq. (36) can be performed by completing the square, and one finds

$$\mathcal{G}^\sigma(\eta,\bar{\eta},s) = \exp\left(\sum_{\ell,\ell'=1}^{L} \bar{\eta}_\ell G^\sigma(\tau_\ell,\tau_{\ell'},s)\eta_{\ell'}\right), \tag{37}$$

with $\tau_\ell = \ell\Delta\tau$, $t_{\ell'} = \ell'\Delta\tau$,

$$G^\sigma(\tau_\ell,\tau_{\ell'},s) = \begin{cases} \frac{I}{I+A_{\ell'}^\sigma(s)}, & \ell = \ell' \ ; \\ B_\ell^\sigma \ldots B_{\ell'+1}^\sigma \frac{I}{I+A_{\ell'}^\sigma(s)}, & \ell > \ell' \ ; \\ -B_{\ell'}^\sigma \ldots B_1^\sigma B_L^\sigma \ldots B_{\ell'+1}^\sigma \frac{I}{I+A_{\ell'}^\sigma(s)}; & \ell < \ell' \ , \end{cases} \tag{38}$$

and

$$A_\ell^\sigma(s) = B_\ell^\sigma \ldots B_1^\sigma B_L^\sigma \ldots B_{\ell+1}^\sigma. \tag{39}$$

By taking appropirate derivatives of Eq. (37) with respect to the components of η_ℓ and $\bar{\eta}_\ell$, we can obtain expressions for any correlation function involving the electron creation and annihilation operators. For example, the single-particle Green's function of Eq. (34) is given by

$$G^\sigma(\tau_\ell,\tau_{\ell'}) = \frac{\sum_s G^\sigma(\tau_\ell,\tau_{ell'},s)\det[I + B_L^+ \ldots B_1^+]\det[I + B_L^- \ldots B_1^-]}{\sum_s \det[I + B_L^+ \ldots B_1^+]\det[I + B_L^- \ldots B_1^-]}. \tag{40}$$

The sums in Eq. (40) are over all Hubbard-Stratonovich spin configurations. In the case of continuous Hubbard-Stratonovich fields, the sums will be replaced by multi-dimensional integrals, and for electron-phonon models, the phonon field will ordinarily replace the Hubbard-Stratonovich field. In the latter cases the product of fermion determinants will be multiplied by $\exp(-S)$, where S depends on the Hubbard-Stratonovich or phonon fields and plays the same role as the Euclidean action in bosonic models. Although we have used the Hubbard model to guide our discussion, it should be clear that the structure of the fermion determinants will be the same for a wide class of models. The only requirement for obtaining Eqs. (33)–(40) is that it be possible to reduce the kinetic and potential energy operators to quadratic forms in the electron creation and annihilation operators.

The Basic Algorithm

We are now in a position to describe the algorithms that we use to carry out fermion simulations. the presence of the fermion determinants makes these algorithms considerably more costly in computer time than those for bosonic systems, because the determinants are highly nonlocal in the Hubbard-Stratonovich fields. In addition, for some models the product of fermion determinants is not positive definite. This introduces very serious difficulties which we shall discuss below. However, we will begin by discussing the simulation of models for which the product of determinants is positive. Examples are the Hubbard model at half-filling and electron-phonon models in which the spin-up and -down electrons couple identically to the phonon field. We shall neglect factors of $\exp(-S)$ in our discussion since the entire difficulty arises from the presence of the fermion determinants. Such factors can be trivially included.

We wish to generate a sequence of Hubbard-Stratonovich configurations, s, with a

probability distribution

$$P(s) = \frac{\det[I + B_L^+ \ldots B_1^+]\, det[I + B_L^- \ldots B_1^-]}{\sum_{s'} \det[I + B_L^+ \ldots B_1^+]\, \det[I + B_L^- \ldots B_1^-]}. \tag{41}$$

The basic idea of importance sampling is to introduce a transition probability, $T_{A\to B}$, for the system to move from configuration s_A to configuration s_B. The only requirements on $T_{A\to B}$ are that detail balance be satisfied,

$$P(s_A) T_{A\to B} = P(s_B) T_{B\to A}, \tag{42}$$

and that it be possible for the system to evolve between any two configurations. Summing Eq. (42) over B and using the fact that $\sum_B T_{A\to B} = 1$ gives

$$P(s_A) = \sum_B P(s_B) T_{B\to A}, \tag{43}$$

so $P(s_B)$ is a left eigenvector of $T_{B\to A}$ with eigenvalue unity. This is the largest eigenvalue, as can be see by considering the general eigenvalue equation

$$\lambda f_A = \sum_B f_B\, T_{B\to A}. \tag{44}$$

Taking the absolute value of both sides of Eq. (44) and then summing over A gives $\lambda \leq 1$. The equality holds only if all of the f_A are positive. This means that starting with any configuration and repeatedly applying $T_{A\to B}$, the system will eventually reach the equilibrium distribution $P(s)$. Once in this distribution, it will remain there.

Clearly the choice of $T_{A\to B}$ is not unique. We need merely require that

$$\frac{T_{A\to B}}{T_{B\to A}} = R^+_{B/A} R^-_{B/A}, \tag{45}$$

where $R^\sigma_{B/A}$ is the ratio of the spin σ fermion determinant in the configuration s_B to that in the configurations s_A. One choice of $T_{A\to B}$ that is particularly convenient for discrete Hubbard-Stratonovich variables is that of the heat bath algorithm,

$$T_{A\to B} = \frac{R^+_{B/A} R^-_{B/A}}{1 + r^+_{B/A} R^-_{B/A}}. \tag{46}$$

Detailed balance is clearly satisfied since $R^\sigma_{A/B} = 1/r^\sigma_{B/A}$.

The procedure is to sweep through the lattice, updating one Hubbard-Stratonovich variable at a time. Whether we use the heat bath form for $T_{A\to B}$ or some other choice, it will be necessary to calculate the determinant ratios $R^\sigma_{B/A}$ in order to decide whether to change each variable. (In the algorithms that we will discuss the configuration s_B will differ from s_A only by the change of one Hubbard-Stratonovich variable.) For a lattice with N spatial sites $I + B^\sigma_L \ldots B^\sigma_1$ is an $N \times N$ dimensional dense matrix. It requires of order N^3 floating-point operators to calculate its determinant. If it were necessary to calculate the determinants in order to update each Hubbard-Stratonovich variable, then of order $N^4 L$ floating-point operations would be required for each sweep of the lattice. This should be compared to the order NL operations we estimated to perform one sweep of the lattice for bosonic simulations.

Fortunately, it is possible to significantly improve the speed of the algorithm by taking into account the structure of the fermion determinant. To be concrete we shall consider the Hubbard model with Hirsch's discrete Hubbard-Stratonovich transformation. However, the extension to other models is straightforward. Suppose that we wish to update the Hubbard-Stratonovich variable on lattice site i and time slice ℓ, $s_{i,\ell}$. Then the final spin configuration s_B differs from the initial configuration s_A only by the substitution $s_{i,\ell} \to -s_{i,\ell}$. First note that the determinants are invariant under a cyclic permutation of the B^σ_ℓ, so

$$\mathcal{Z}_\sigma(s_A) = \det[I + B^\sigma_\ell \ldots B^\sigma_1 B^\sigma_L \ldots B^\sigma_{\ell+1}] = \det[I + A^\sigma_\ell(s_A)]. \tag{47}$$

Since the matrix u^σ_ℓ is diagonal and has only one element which depends on $s_{i,\ell}$, we see that

$$A^\sigma_\ell(s_B) = [I + \Delta^\sigma(i,\ell)] A^\sigma_\ell(s_A), \tag{48}$$

where $\Delta^\sigma(i,\ell)$ is a matrix with only one nonzero element,

$$\Delta^\pm(i,\ell)_{jk} = \delta_{ji}\delta_{ki}\left(e^{\pm 2\Delta\tau\lambda s_{i,\ell}} - 1\right). \tag{49}$$

¿From Eq. (48) we see that the ratio of determinants for spin configurations s_B and s_A is

$$R^{\sigma}_{B/A} = \frac{\det[I + A^{\sigma}_{\ell}(s_B)]}{\det[I + A^{\sigma}_{\ell}(s_A)]}$$
$$= \det[I + \Delta^{\sigma}(i,\ell) A^{\sigma}_{\ell}(s_A) G^{\sigma}(\tau_\ell, \tau_\ell, s_A)]. \tag{50}$$

Notice that the quantity in square brackets in the last line of Eq. (50) is the sum of the unit matrix and a matrix with only one nonzero row. The determinant can therefore be evaluated immediately, and we find

$$R^{\sigma}_{B/A} = 1 + [1 - G^{\sigma}(\tau_\ell, k\tau_\ell, s_A)_{ii}]\Delta^{\sigma}(i,\ell)_{ii}. \tag{51}$$

Thus, the ratios of determinants needed for the importance sampling decision can be obtained with very few floating-point operations provided the equal-time Green's functions are known. Notice that the locality of the original fermion interaction is essential for this result.

If the equal-time Green's functions in the configurations s_A are known, and the flip of the spin $s_{i,\ell}$ is accepted, then the Green's functions in the configuration s_B can be obtained from the identity

$$G^{\sigma}(\tau_\ell, \tau_\ell, s_B) = G^{\sigma}(\tau_\ell, \tau_\ell, s_A) - G^{\sigma}(\tau_\ell, \tau_\ell, s_A)\Delta^{\sigma}(i,\ell) A^{\sigma}_{\ell}(s_A) G^{\sigma}(\tau_\ell, \tau_\ell, s_B). \tag{52}$$

The fact that $\Delta^{\sigma}(i,\ell)$ has only one nonzero matrix element means that Eq. (52) and be solved immediately, yielding

$$G^{\sigma}(\tau_\ell, \tau_\ell, s_B) = G^{\sigma}(\tau_\ell, \tau_\ell, s_A) - \frac{G^{\sigma}(\tau_\ell, \tau_\ell, s_A)\Delta^{\sigma}(i,\ell)[I - G^{\sigma}(\tau_\ell, \tau_\ell, s_A)]}{1 + [1 - G^{\sigma}(\tau_\ell, \tau_\ell, s_A)_{ii}]\Delta^{\sigma}(i,\ell)_{ii}}. \tag{53}$$

There are no matrix multiplications in Eq. (53), so of order N^2 floating points operations are required to update the equal time Green's functions. As a result, of order $N^3 L$ operations are required to update all of the Hubbard-Stratonovich variables once. This represents a considerable savings in computer time over evaluating the fermion determinants at each

step in the calculation, but is still significantly more costly than the corresponding calculation with boson degrees of freedom. For models with off-site electron interactions, it will be necessary to evaluate a low-dimension determinant to evaluate $R_{B/A}$, and to solve a corresponding set of linear equations to update the equal-time Green's functions. However, the order of magnitude of the loating point operations will not change as long as the interactions are short range.

To begin the calculation it is necessary to form $A^\sigma(\ell)$ for one time slice by multiplying the B_ℓ^σ matrices together according to Eq. (39). One can then form the corresponding equal-time Green's functions by inverting the matrices $I + A^\sigma(\ell)$ using, or example, Gaussian elimination. This procedure must be repeated periodically during the simulation in order to avoid the buildup of roundoff errors. Thus we need to be able to multiply the B_ℓ^σ rapidly. This is also necessary in moving between time slices. After all the Hubbard-Stratonovich variables on a given time slice have been updated, we move on to the next one, using the relation

$$G^\sigma(\tau_{\ell+1}, \tau_{\ell+1}, s) = B_{\ell+1}^\sigma G^\sigma(\tau_\ell, \tau_\ell, s) B_{\ell+1}^\sigma - 1$$

to obtain the equal-time Green's functions. for the Hubbard model u_ℓ^σ is a diagonal matrix, so there is no problem in forming and multiplying by $\exp(-\Delta\tau u_\ell^\sigma)$. Although the hopping matrix, k, is sparse, $\exp(-\Delta\tau k)$ is not. One possibility is to simply expand the exponential to order $\Delta\tau^2$. The approach we generally use is to introduce further Trotter breakups so that we only deal with sparse matrices. For example, for a two-dimensional problem we can write $k = k_x + k_y$, where k_x and k_y involve hopping in the x and y directions respectively. We further subdivide these matrices so that $k_x = k_x^o + k_x^e$, where k_x^o contains hopping between lattice sites (1,2),(3,4),... in the x-direction, and k_x^e contains hopping between sites (2,3),(4,5)... in the x-direction with a similar breakup for k_y. We can then write

$$e^{-\Delta\tau k} \simeq e^{-\Delta\tau k_x^o} e^{-\Delta\tau k_x^e} e^{-\Delta\tau k_y^o} e^{-\Delta\tau k_y^e}. \tag{55}$$

Since the individual terms in $k_x^o \ldots$ decouple, the right-hand side of Eq. (55) is a product of sparse matrices. This additional Trotter breakup introduces errors of order $\Delta\tau^2$ in measured quantities, which is the same order of magnitude as the errors arising from the breakup of Eq. (3). The same approach can be used for evaluating $\exp(-\Delta\tau u_\ell^\sigma)$ for models in which the change in a single Hubbard-Stratonovich variable alters more than one element of u_ℓ^σ. If we can write A_ℓ^σ as a product of sparse matrices, then it will require of order N^2L floating point operations to evaluate it. Of order N^3 operations are needed to perform the matrix inverse required to obtain the equal-time Green's functions. Thus, as long as we can update a few time slices between recalculation of the Green's functions, this part of the algorithm will not dominate the computer time. The total time for updating all Hubbard-Stratonovich variables will grow with N and L as N^3L.

Low-Temperature Numerical Instabilities

We have now assembled all the ingredients for the original form of the algorithm. Indeed, the algorithm we have just described has been used to carry out studies of a variety of many-electron models. However, numerical instabilities have prevented its applications at low temperatures. To perform simulations at low temperatures one must identify the cause of these instabilities and remove it. The instabilities are not associated with the updating of the equal-time Green's functions by Eq. (53). Indeed, we have found that by staying on the same time slice, one can update tens of thousands of Hubbard-Stratonovich variables without accumulating roundoff errors. Numerical instabilities to arise from advancing the Green's functions to new time slices via Eq. (54), and most importantly in the construction of A_ℓ^σ itself. The origin of the problem can already be seen in the free theory. In d spatial dimensions the eigenvalues of A_ℓ^σ range between $\exp(2\,dt\beta)$ and $\exp(-2\,dt\beta)$, so this matrix becomes very badly conditioned as the temperature is lowered. The situation only

worsens for interacting theories. The eigenvalues of A_ℓ^σ that are physically important are typically of order unity, and it is necessary to computer A_ℓ^σ accurately enough to incorporate information about them. It is clear that we must separate the contributions of the large and small eigenvalues or the latter will be completely swamped by roundoff errors of the former. This point was first made by Sugiyama and Koonin in their development of an algorithm for the calculation of ground-state properties. They made use of an orthogonalization procedure to separate the diverse energy scales in the fermion matrices. We have followed a similar approach based on matrix factorization in our finite-temperature calculations.

Suppose that one can multiply m of the B_ℓ^σ without losing numerical control. We then use the Gram-Schmidt orthogonalization procedure to write this product in the form

$$a_1^\sigma(\ell) = B_{\ell+m}^\sigma B_{\ell+m-1}^\sigma \cdots B_{\ell+1}^\sigma = U_1^\sigma D_1^\sigma R_1^\sigma. \tag{56}$$

Here U_1^σ is an orthogonal matrix, D_1^σ a diagonal matrix, and R_1^σ a right triangular matrix with diagonal elements equal to unity. the orthogonal matrix U_1^σ is necessarily well conditioned; *a priori*, R_1^σ need not be well conditioned, but in practice we find that it is. Only the diagonal matrix D_1^σ has large variations in the size of its elements.

We next form

$$\begin{aligned} a_2^\sigma(\ell) &= B_{\ell+2m}^\sigma \cdots B_{\ell+1}^\sigma \\ &= B_{\ell+2m}^\sigma \cdots B_{\ell+m+1}^\sigma U_1^\sigma D_1^\sigma R_1^\sigma \\ &= U_2^\sigma D_2^\sigma R_2^\sigma. \end{aligned} \tag{57}$$

The order of operations in Eq. (57) is important. We first multiply U_1^σ by $B_{\ell+2m}^\sigma \cdots B_{\ell+m+1}^\sigma$. By assumption m is small enough so that this matrix can be computed accurately. We then multiply it on the right by D_1^σ. This only rescales the columns of the matrix, and thus does no harm to the numerical stability of the next step, a UDR decomposition of this partial product. We then multiply the resulting triangular matrix on the right by R_1^σ to obtain

the last line of Eq. (57). This process is repeated L/m times to obtain

$$a_\ell^\sigma(s) = a_{L/m}^\sigma(\ell) = U_{L/m}^\sigma d_{L/m}^\sigma R_{L/m}^\sigma. \tag{58}$$

In forming $I + A_\ell^\sigma, (s)$, care must be taken to isolate the diagonal matrix, $D_{L/m}^\sigma$, whose elements have large variations in size. We therefore write

$$I + A^\sigma(\ell) = U_{L/m}^\sigma \left[U_{L/m}^{\sigma -1} R_{L/m}^{\sigma -1} + D_{L/m}^\sigma \right] R_{L/m}^\sigma. \tag{59}$$

The matrices $U_{L/m}^\sigma$ and $R_{L/m}^\sigma$ are, of course, trivial to invert. We can form the inverse of the matrix in square brackets by Gaussian elimination. This procedure is numerically stable because the large and small eigenvalues are again separated. Alternatively, we can form a last UDR decomposition of the matrix in square brackets, which is also numerically stable. We have tested the numerical stability of our method for evaluating $G^\sigma(\tau_\ell, \tau_\ell, s)$ for values of $t\beta$ as large as 100 without encountering problems.

The Fermion Sign Problem

Up to now we have assumed that the product of fermion determinants for the spin up and down electrons is positive semi-definite, so that $P(s)$ as defined in Eq. (41) can be interpreted as a probability. This will be the case for electron-phonon models in which the electrons in the two spin states have identical couplings to the phonon field. Then the two determinants are equal real numbers, so their product must be positive semi-definite. the same is true for the Hubbard model with an attractive Coulomb interaction, $U < 0$, since in this model the spin-up and -down electrons couple identically to the Hubbard-Stratonovich field. for the Hubbard model with a repulsive Coulomb interaction, $U > 0$, the spin-up and spin-down electrons couple to the Hubbard Stratonovich field with opposite signs, as we have seen in Eqs. (19) and (21), so the determinants are not equal. However, their product

is positive semi-definite if we set $\mu = 0$ in Eq. (9). To see this we make a particle-hole transformation on the spin-down electrons, by introducing the unitary operator $U_{\rm ph}^-$,

$$\begin{aligned} U_{\rm ph}^- c_{i-} U_{\rm ph}^{-\dagger} &= (-)^i c_{i-}^\dagger \\ U_{\rm ph}^- n_{i-} U_{\rm ph}^{-\dagger} &= 1 - n_{i-}. \end{aligned} \tag{60}$$

Now $c_-^\dagger \cdot k \cdot c_-$ is invariant under this transformation, while $c_-^\dagger \cdot u_\ell^- \cdot c_-$ changes sign, so

$$\begin{aligned} \mathcal{Z}_-(s) &= \mathrm{Tr}\left(U_{\rm ph}^- \, e^{-\Delta\tau c_-^\dagger \cdot u_L^- \cdot c_-} \ldots e^{-\Delta\tau c_-^\dagger \cdot k \cdot c_-} U_{\rm ph}^{-\dagger} \right) \\ &= e^{\Delta\tau\lambda \sum_{i,\ell} s_{i,\ell}} \, \mathcal{Z}_+(s). \end{aligned} \tag{61}$$

Thus, the spin-up and spin-down fermion determinants are proportional to each other with a positive constant of proportionality, so the product of determinants is again positive. Note that this argument does not hold for $\mu \neq 0$. Finally, note that for $\mu = 0$ the Hamiltonian of Eq. (9) is invariant under a particle-hole transformation on both the spin-up and spin-down electrons. Thus, for this case $\langle n_{i+} + n_{i-} \rangle = \langle 1 - n_{i+} - n_{i-} \rangle$, so for $\mu = 0$ there is an average of one electron per site, a half-filled band.

For the examples just cited, the probability distribution of Eq. (41) can be used in simulations. However, it should be kept in mind that the individual fermion determinants can, and do, change sign. This happens with increasing frequency as the temperature is lowered and/or the coupling is increased. One must verify that the algorithm being used allows the system to cross surfaces in the Hubbard-Stratonovich or phonon variables along which the determinants vanish. Otherwise, the entire space will not be explored. It has been verified with some care that this is the case for the Monte Carlo algorithm described above for both discrete and continuous Hubbard-Stratonovich variables. The important point is that at each step in the algorithm one makes a finite change in one of the variables, which can cause the system to cross a surface of vanishing determinant. By contrast, great care must be taken with algorithms based on the Langevin or molecular dynamics equations, in which all variables are changed by an infinitesimal amount at each step. The surfaces along

which the determinants vanish correspond to infinite potential barriers for these equations, and can only be crossed through errors in their numerical integration. Even if the barriers are crossed, the system is generally thrown out of statistical equilibrium.

For models which lack a symmetry relating the spin-up and -down fermion determinants, the product of determinants will generally fluctuate in sign. For example, this is the case for the Hubbard model with repulsive Coulomb interaction for $\mu \neq 0$. Our approach must then be slightly modified. Let us first consider the partition function

$$Z = \sum_s \mathcal{Z}_+(s)\mathcal{Z}_-(s), \tag{62}$$

with $\mathcal{Z}_\sigma(s)$ given by Eq. (47). We can always write the product of fermion determinants in the form

$$\mathcal{Z}_+(x)\ddagger_-(s) = \tilde{P}(s)\int(s), \tag{63}$$

where $\tilde{P}(s)$ is a positive semi-definite probability function normalized so that

$$\sum_s \tilde{P}(s) = 1. \tag{64}$$

If we choose Hubbard-Stratonovich configurations with the probability $\tilde{P}(s)$, then the partition function is given by the average value of $\mathcal{S}(s)$ in this distribution,

$$Z = \langle \mathcal{S} \rangle_{\tilde{P}}. \tag{65}$$

Here and below we denote averages with respect to the distribution $\tilde{P}(s)$ by $\langle \rangle_{\tilde{P}}$. Of course Eq. (65) holds for any choice of $\tilde{P}(s)$ provided Eqs. (63) and (64) are satisfied. We determine $\tilde{P}(s)$ by minimizing the variance in the calculation of Z. That is, we find the minimum of $\langle ^2 \rangle_{\tilde{P}}$ as a function of the $\tilde{P}(s)$, subject to the constraints of Eqs. (63) and (64). (Of course $\langle \mathcal{S} \rangle_{\tilde{P}}$ is invariant under changes in $\tilde{P}(s)$ as long as Eq. (63) is satisfied.) a short calculation yields

$$\tilde{P}(s) = \frac{|Z_+(s)Z_-(s)|}{\sum_{s'} |Z_+(s')Z_-(s')|} \tag{66}$$

smallskip

$$\mathcal{S}(s) = S(s) \sum_{s'} |Z_+(s')Z_-(s')|. \tag{67}$$