Fermion sign problem: Decoupling transformation and simulation algorithm

Ghassan George Batrouni*
Thinking Machines Corporation, 245 First Street, Cambridge, Massachusetts 02142

Philippe de Forcrand†
Theoretical Physics Institute, University of Minnesota, Minneapolis, Minnesota 55455

(Received 18 November 1992; revised manuscript received 6 April 1993)

We discuss the fermion sign problem and, by examining a very general Hubbard-Stratonovich transformation, argue that the sign problem cannot be solved with such methods. We propose a different kind of transformation which, while not solving the sign problem, shows more detailed information about the system. With our transformation it is *trivial* to tell which auxiliary field configurations give a positive sign and which give a negative sign. We then discuss briefly various properties of this transformation and construct an algorithm which with one simulation gives results for a whole range of particle densities and Hubbard U values, positive and negative. Our approach is in excellent agreement with exact calculations.

The major obstacle facing numerical simulation of a large number of strongly interacting electrons is the so-called "fermion sign problem."[1-7] This problem appears in many different guises and here we will focus on its form in the determinant algorithm for quantum Monte Carlo. Our method and conclusions are general but for clarity we will concentrate on the Hubbard model, which is under active study as a model for metal-insulator transitions and high-temperature superconductivity. We begin by reviewing the Hubbard Stratonovich transformation and the resulting sign problem. The partition function of the Hubbard model in the grand canonical ensemble is given by

$$Z=\mathrm{tr}(e^{-\beta H}) , \tag{1}$$

$$H=-t\sum_{\langle ij\rangle,\sigma}[c_\sigma^\dagger(i)c_\sigma(j)+c_\sigma^\dagger(j)c_\sigma(i)]+U\sum_i[n_+(i)-\tfrac{1}{2}][n_-(i)-\tfrac{1}{2}]$$
$$-\mu\sum_i[n_+(i)+n_-(i)] , \tag{2a}$$

$$=\sum_{ij,\sigma}c_\sigma^\dagger(i)k_{ij}c_\sigma(j)+U\sum_i[n_+(i)-\tfrac{1}{2}][n_-(i)-\tfrac{1}{2}] . \tag{2b}$$

The sum $\langle ij\rangle$ is over all pairs of nearest neighbor lattice sites, t is the hopping parameter, $c_\sigma^\dagger(i)$ and $c_\sigma(i)$ are creation and annihilation operators of electrons with spin σ along the z axis at site i. n_+ (n_-) is the number operator for electrons with up (down) spins. β is the inverse temperature, and U is the coupling constant, which can be positive or negative. The matrix k in Eq. (2b) contains the hopping term and the chemical potential. Through standard techniques one can transform the trace into a path integral (or sum) over the configurations of a c-number auxiliary field.[8] First we use the Trotter-Suzuki approximation[9] to express Z as $Z=\mathrm{tr}(e^{-\tau H})^L \approx \mathrm{tr}(e^{-\tau k}e^{-\tau V})^L$, where $\tau=\beta/L\ll 1$, is the imaginary time step, and L is the number of such time steps. Now we use the Hubbard-Stratonovich (HS) transformation to decouple the quartic potential term $e^{-\tau V}$ into quadratics in the creation and annihilation operators:

$$\exp\{-\tau U[n_+(i)-\tfrac{1}{2}][n_-(i)-\tfrac{1}{2}]\}=\frac{\exp(-\tau U/4)}{2}\sum_{s(i,l)=\pm 1}\exp\{-\lambda s(i,l)[n_+(i)-n_-(i)]\} \tag{3}$$

at each site i and time slice l. λ is related to U via $\cosh(\lambda)=e^{\tau U/2}$. Here we write the discrete HS transformation,[4] but one can also write a continuous one. The trace over the fermion operators can now be taken, since they only appear quadratically. This gives[8]

$$Z=\sum_{s(i,l)=\pm 1}\det M^+\det M^- , \tag{4}$$

where

$$M^\sigma=I+B_L^\sigma B_{L-1}^\sigma\cdots B_1^\sigma \tag{5}$$

and

$$B_l^\pm=e^{\mp\lambda v(l)}e^{-\tau k} . \tag{6}$$

I is a $V\times V$ unit matrix, V is the spatial volume,

$v(l)_{ij}=\delta_{ij}s(i,l)$, where i runs from 1 to V and l from 1 to L. The partition function is now written as a sum over c numbers and can therefore be simulated on a computer. The sign problem arises because the determinants and their product, in Eq. (4), can be negative and thus cannot be used as the probability density in a Monte Carlo simulation. Instead, their absolute value is used, and the average of an observable A is then given by $\langle A\rangle=\langle A\,\mathrm{sgn}\rangle'/\langle\mathrm{sgn}\rangle'$, where $\langle\ \rangle'$ denotes averages with respect to the absolute value of the determinants. The average sign is thus defined as $\langle\mathrm{sgn}\rangle'=Z/Z'$, where Z' is the partition function resulting from using the absolute value of the determinants. This approach works well for small U and relatively high temperatures. As the temperature decreases (β increases), the average sign scales like[3,5] $\langle\mathrm{sgn}\rangle\sim e^{-c\beta}$, where c is a constant. This makes low-temperature simulations impractical because averages are obtained as the ratios of two very small numbers, each with a very large variance. When $U<0$, a similar procedure yields $M^{+}=M^{-}$. Consequently, the product of the two determinants in Eq. (4) becomes a square, and therefore positive semidefinite even though the determinant itself is still *not* positive semidefinite. Thus, there is no sign problem for the negative-U Hubbard model (or other attractive interactions).

The HS transformation discussed above is the most commonly used one, but it is only one of an infinite number of possible transformations. For example, other decoupling schemes were discussed in Refs. 5 and 6 in the hope of finding a transformation, which will solve the sign problem or at least decrease its severity. We will argue here that there exists no HS transformation that will eliminate the sign problem. We start by noting that the purpose of any HS transformation is to decouple the quartic fermionic interaction into quadratic terms which are coupled to the HS (auxiliary) field. This allows us to perform the trace in the partition function, giving two determinants. But, except for minor details, all transformations examined so far have yielded either a product of different determinants of the above form, or a single determinant,[5] which have always suffered severely from the sign problem for certain values of μ, β, and U. This suggests that one way to solve the sign problem in this approach is to obtain a square of a determinant. Is it therefore possible to generalize the above HS transformations such that the resulting partition function is a sum (or integral) over a square, and if not, why? In order to get $(\det M)^2$, the relative minus sign between n_- and n_+ on the right-hand side of Eq. (3) must become a plus, thus preserving the up-down symmetry of the Hamiltonian in the HS transformation. Let us therefore propose a general HS transformation

$$\exp[-\tau U(n_+n_- -\tfrac{1}{2}n_+ -\tfrac{1}{2}n_-)]=\int dy P(y)y^{(n_++n_-)}\ , \tag{7}$$

where $P(y)$ and y are real[10] and arbitrary except for the constraints discussed below, and U is positive or negative. This transformation includes discrete transformations like Eq. (3), but is more general. If such a transformation were possible, the sign problem would be solved because of the resulting $(\det M)^2$. Since $n_\pm=0,1$, we see that the conditions on $P(y)$ and y are

$$\int dy\, P(y)=1\ , \tag{8}$$

$$\langle y\rangle=e^{\tau U/2}\ , \tag{9}$$

$$\langle y^2\rangle=1\ , \tag{10}$$

where $\langle\ \rangle$ means an average with respect to the weight $P(y)$. In general, the inequality $\langle y\rangle^2\le\langle y^2\rangle$ must be satisfied. Combining this with Eqs. (9) and (10) forces U to be negative: In other words, the inequality cannot be satisfied for positive values of U. Therefore, there is no general HS transformation that is capable of giving the square of a determinant. Consequently, this approach to solving the sign problem fails.

Implicit in the above argument is the positivity of $P(y)$. We will reserve the name "Hubbard-Stratonovich transformation" for these cases, since all previous applications of the HS transformation assume such positivity. However, this argument is invalid if we allow $P(y)$ to take negative values. This of course would *not* solve the sign problem, but it would give us a square of a determinant, with the result that the sign changes now come not from the determinants, but from $P(y)$. One advantage of this is that contrary to the HS transformation, where, in general, we do not know which auxiliary field configurations lead to minus signs because the structure of the determinants is too complicated; here the minus sign comes from $P(y)$, which we know exactly. Therefore, we have complete prior knowledge of the sign of all the configurations.

For example, a simple choice is

$$P(y)=\prod_{i,l}[a\delta(y(i,l)-y_1)+b\delta(y(i,l)-y_2)]\ , \tag{11}$$

where i is the space index, and l is the time slice index. y_1 and y_2 are the allowed values for the auxiliary field y, a, and b are parameters to be determined from the constraints Eqs. (8), (9), and (10). We chose two discrete values, y_1 and y_2, for the auxiliary field, but we could have chosen any number of discrete variables, or any continuous distributions as long as we satisfy conditions Eqs. (8), (9), and (10). Our motivation is simplicity. Applying the above constraints gives $a+b=1$ and

$$b=\frac{y_1^2-1}{y_1^2-y_2^2}\ , \tag{12}$$

$$e^{\tau U/2}=\frac{1+y_1y_2}{y_1+y_2}\ . \tag{13}$$

Having chosen the form of $P(y)$, we still have the freedom of choosing the values of one of the two parameters y_1,y_2, the second being determined by Eq. (13).

Now that we have decoupled the quartic fermion interaction into two quadratic terms with the same sign, the trace over the Fermi operators can be performed as in the HS case giving for the partition function

$$Z=\sum_{y(i,l)=y_1,y_2} a^{n_1}b^{n_2}(\det M)^2 \tag{14}$$

$$=a^N\sum_{n_2} C^N_{n_2}\left[\frac{b}{a}\right]^{n_2}\langle(\det M)^2\rangle_{n_2}\ . \tag{15}$$

M has the same form as M^-, Eq. (5), and

$$B_l=v(l)e^{-\tau k}\ , \tag{16}$$

with $v(l)_{i,j}=\delta_{ij}y(i,l)$. n_1 (n_2) is the number of auxiliary spins with the value y_1 (y_2), and $N=n_1+n_2$ is the total number of sites on the $(d+1)$ dimensional lattice. $\langle\ \rangle_{n_2}$ is an average over all configurations which have n_2 spins equal to y_2, whose number is the binomial coefficient $C^N_{n_2}$. It is easy to show that for $U<0$, both a and b are positive and therefore there is no sign problem, just like the usual HS transformations. When $U>0$, a and b have *opposite* signs (we take $b<0$), and thus the sign problem reappears. However, although the value of $(\det M)^2$ (always positive) depends on *both* the relative number of spins with values y_1 and y_2, and their configuration on the lattice, the prefactor $a^{n_1}b^{n_2}$ depends *only* on the relative numbers. In particular, since the source of the sign problem in our formulation is the opposite sign of a and b, we have complete knowledge of all the configurations that change the sign: only configurations with an odd number of y_2 spins lead to an odd exponent for b and thus a negative contribution. This complete characterization of the negative configurations is to be contrasted with all the HS transformations previously used where one knows very little about the kind of configurations that lead to minus signs. This vividly demonstrates, yet again,[5] that the sign problem in the determinant algorithm is not related to configurations where the electron "paths" exchange. It is merely an artifact of the transformation used to decouple the quartic terms in the Hamiltonian. Another property of our transformation is that it preserves the rotational spin symmetry of the original Hamiltonian. We can therefore take measurements along any spin direction, or along all three, thus reducing fluctuations.

Not surprisingly, the behavior of the average sign and the usual summand in Eqs. (14) and (15) depend on the values we choose for the auxiliary fields. Three choices are of particular interest to us. The first is to choose $y_1\approx y_2$. This choice is remarkable because the entire phase space is explored with a variable that fluctuates very little. Consequently the values of the determinant and other observables fluctuate little, and change very smoothly as more spins are flipped. The disadvantage of this choice is that $a/b\to-1$ (always keeping $a+b=1$) as $y_1\to y_2$, which means that the contributions of the negative configurations are of the same size as the positive ones. This makes numerical simulations hard, but may offer the possibility of studying the system semi-analytically.

Our second choice is to take $y_2=0$. What is intriguing about this choice is that when enough auxiliary spins have the value $y_2=0$, there will be many realizations where the auxiliary field for at least one entire time slice is zero, preventing the percolation of y_1 spins and resulting in $M=I$ and $\det M=1$. When this happens the observables will also have a trivial value. So, when such configurations become important, there will be many instances where the observables are trivial and it would be interesting to study the effect such configurations might have on phase transitions.

The third choice is to take $y=\pm1$. The limit of Eqs. (12) and (13) as $y_1\to+1$ and $y_2\to-1$ is

$$a=\tfrac{1}{2}(1+e^{\tau U/2})\ , \tag{17}$$

$$b=\tfrac{1}{2}(1-e^{\tau U/2})\ . \tag{18}$$

Notice that with this choice, M is *no longer* a function of the coupling constant U, since the coupling constant can appear only in y, a, and b, and here we fixed $y=\pm1$. Furthermore, recall that, in the matrix M, μ appears in the form $e^{\tau\mu}I$ for each B^σ_l matrix, where I is the $V\times V$ identity. Consequently, the matrix $M'=e^{-\beta\mu}(M-I)$ is a function of temperature, but *not* U or μ. This major simplification allows us to collect a large number of realizations of M', and then perform data analysis for any U and μ. The μ dependence is obtained by multiplying M' by $e^{\beta\mu}$. The U dependence appears through the ratio b/a, Eqs. (15), (17), and (18), in a way similar to perturbation theory for τ small. We can thus obtain results for a wide range of values of U (positive and negative) and μ from only one computer run. Currently, we generate the realizations of M' as follows: start with all $y=+1$ and successively flip randomly selected spins to $y=-1$, calculating M' after each spin flip. This is done up to some maximum number of flipped spins, much smaller than the total number of sites. The reason is that the expansion of Z in n_2 [Eq. (15)], which is similar to a perturbation expansion for τ small, converges quickly and can be truncated. Then we start all over, and in this way we generate an ensemble of realizations of flipped spins over which we can average. The addition of importance sampling could greatly increase the efficiency of the algorithm. This is currently being investigated.

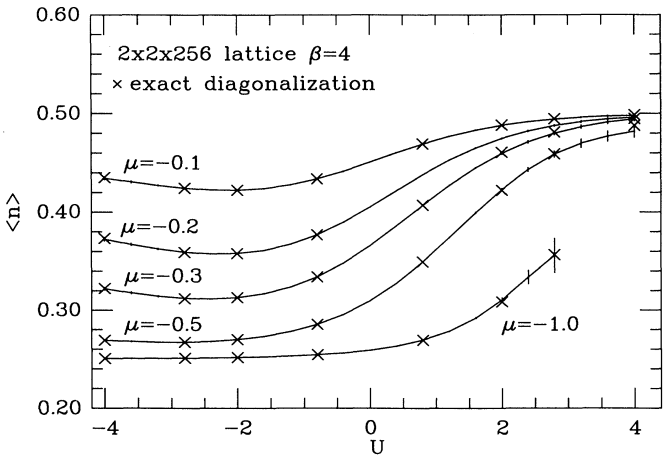

FIG. 1. The average density $\langle n_\uparrow\rangle$ vs U for $\mu=-0.1$, -0.2, -0.3, -0.5, and -1 as labeled in the figure. The crosses show exact results.

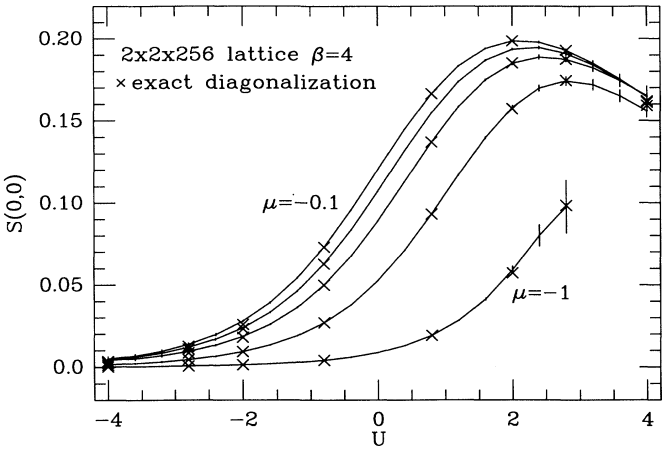

FIG. 2. The ferromagnetic correlation function $S(0,0)$ vs U for the same values of the chemical potential as in Fig. 1. The crosses show exact results.

We tested this algorithm on a 2×2 lattice and compared with exact results. In Fig. 1 we show a plot of $\langle n_\uparrow \rangle$ versus U, with the crosses showing exact results. Figure 2 shows a similar figure for the ferromagnetic correlation function, $S(0,0)$. We used 256 time slices in order to eliminate the finite time-step errors for comparison with exact diagonalization. In general, the finite τ errors are $O(\tau^2 U)$, as in the usual determinant algorithm. The major part of the computing effort goes into the evaluation of the determinant, which scales like the cube of the volume of the system. Note that as we move away from half filling the errors increase appreciably for positive U because the sign problem is appearing in force. Just like in the usual case, the average sign decreases exponentially with the inverse temperature, and therefore the computer time required to keep statistical errors under control also grows exponentially. It is worth noting, however, that here we can reduce the fluctuations of the average sign by an order of magnitude or more, by grouping together positive and negative configurations:[7] This is possible in our formulation only, because of our *a priori* knowledge of the sign. For $U<0$, where there is no sign problem, errors are very small. Recall that all of the shown curves were obtained from the data of only one run, and that those data contain all the information needed to measure all the equal time correlation functions for positive and negative U and a wide range of μ.

We would like to thank A. Krasnitz and R. Scalettar for very helpful discussions and (R.S.) for providing exact diagonalization results against which we compared our algorithm. G.G.B. thanks the Theoretical Physics Institute at the University of Minnesota for its hospitality when this work was in its early stages.

*Electronic address: ggb@think.com
†Permanent address: IPS, Eidgenössische Technische Hochschule Zürich, CH-8092 Zurich, Switzerland. Electronic address: forcrand@ips.id.ethz.ch

[1]E. Y. Loh, Jr., J. E. Gubernatis, R. T. Scaletter, S. R. White, D. J. Scalapino, and R. L. Sugar, Phys. Rev. B **41**, 9301 (1990).
[2]S. Sorella, S. Baroni, R. Car, and M. Parinello, Europhys. Lett. **8**, 663 (1989); S. Sorella, E. Tosatti, S. Baroni, R. Car, and M. Parinello, Int. J. Mod. Phys. B **1**, 993 (1988).
[3]S. R. White, D. J. Scalapino, R. L. Sugar, E. Y. Loh, J. E. Gubernatis, and R. T. Scalettar, Phys. Rev. B **40**, 506 (1989).
[4]J. E. Hirsch, Phys. Rev. B **31**, 4403 (1985).
[5]G. G. Batrouni and R. T. Scalettar, Phys. Rev. B **42**, 2282 (1990).
[6]Liang Chen and A.-M.S. Tremblay, Int. J. Mod. Phys. B **6**, 547 (1992).
[7]F. F. Assaad and Ph. de Forcrand, in *Quantum Simulations of Condensed Matter Phenomena* (World Scientific, Singapore, 1990), pp. 1–11.
[8]R. Blankenbecler, D. J. Scalapino, and R. L. Sugar, Phys. Rev. D **24**, 2278 (1981).
[9]H. F. Trotter, Proc. Am. Math. Soc. **10**, 545 (1959); M. Suzuki, Commun. Math. Phys. **51**, 183 (1976).
[10]This can be easily generalized to complex variables and $\det(M^\dagger M)$, but to keep the notation simple we use real values.