PHYSICAL REVIEW E VOLUME 50, NUMBER 4 OCTOBER 1994

Transforming signs to phase distributions in quantum simulations

J. M. Deutsch
University of California, Santa Cruz, Santa Cruz, California 95064
(Received 20 April 1994)

A method is developed which speeds up averaging in quantum simulations where minus signs cause difficulties. A Langevin equation method in conjunction with a replication algorithm is used, enabling one to average over a continuously varying complex number. Instead of ensemble averaging this number directly, the phase of the complex number is followed over time. The method is illustrated in some simple cases where the answers obtained can be compared to exact results, and also compared to conventional averaging procedures which converge orders of magnitude more slowly than this method. Limitations of this method are also described.

PACS number(s): 05.30.−d

I. INTRODUCTION

Simulation of quantum many body systems have been greatly hindered by the "sign problem." Quantum systems with many degrees of freedom are often simulated using stochastic methods. For a variety of problems, usually involving fermions, an average over many negative and positive numbers must be taken. Sometimes the sign does not affect results [1] but in many interesting cases the resulting average is very small because of cancellations, so that in order to obtain good statistics, the number of realizations that must be taken is exponential in the inverse temperature, making it infeasible to get information about ground state properties of a system.

The main idea of this work can be understood by considering the following example. It will be seen below that it describes the quantum mechanics of n bosons restricted to one site. Consider the stochastic equation for $z(t)$ which is a complex function of time:

$$\dot{z}=if(t)z. \tag{1}$$

Here $f(t)$ is real Gaussian noise with correlation function $\langle f(t)f(t')\rangle = v\delta(t-t')$, where the angle brackets denote an average over the noise $f(t)$. Choosing $z(0)=1$ the above equation is easily solved giving

$$z=e^{i\theta(t)}, \tag{2}$$

where

$$\theta(t) = \int_0^t f(\tau)d\tau. \tag{3}$$

Here $\theta(t)$ describes a random walk in time, and therefore the probability distribution of θ at time t, $P(\theta,t)$, is a Gaussian distribution with zero mean and variance $\sigma^2 = vt$. Below we will see that the ground state energy of n bosons is simply obtained from the asymptotic behavior of $\langle z(t)^n\rangle$, which is solved for as follows:

$$\langle z(t)^n\rangle = \langle e^{in\theta}\rangle = \int P(\theta,t)e^{in\theta}d\theta = e^{-n^2vt/2}. \tag{4}$$

Now let us turn to the problem of how, without the aid of this analytical solution, one could obtain $\langle z(t)^n\rangle$ numerically. Because $|z(t)| = 1$, to obtain the correct average would require over e^{n^2vt} independent runs of Eq. (1) to obtain adequate statistics. This can be very large.

The alternative to this direct averaging procedure is to obtain the probability distribution $P(\theta,t)$. In this case it is much better to keep track of the total angle θ, not $\theta \bmod(2\pi)$. Keeping track of the total angle is done as Eq. (1) is evolved. The increment in θ found in a single time step is added to the previous value of θ, and a histogram of θ at a given time can be obtained by running Eq. (1) over many realizations of f.

Does analytically continuing θ and then finding $P(\theta,t)$ help in obtaining $\langle z(t)^n\rangle$? In this case and in the less trivial cases considered below, it appears to reduce the computation time from something exponential to algebraic. Of course assumptions must be made in order to obtain such a reduction, the most important one being that $P(\theta,t)$ is a smooth function of θ. There are also problems encountered in deciding what function to use in fitting P, as will be discussed later. In the case of interest here, fitting $\ln P(\theta,t)$ to a quadratic function of θ gives accurate results. In fact, the ground state energies should be obtained to within an error $N^{-1/2}$ where N is the number of independent runs, in contrast to direct averaging which gives an error of order $\exp(n^2vt)N^{-1/2}$.

Having motivated the method that will be considered, the Langevin equation that simulates quantum many body systems will be described and numerical results will be given.

II. SIMULATION METHOD

A system of fermions or bosons can be simulated by considering a stochastic equation which is a generalization of Eq. (1). A field $\phi(\mathbf{r},t)$ is governed by the equation

$$\dot{\phi}(\mathbf{r},t)=f(\mathbf{r},t)\phi+\frac{1}{2m}\nabla^2\phi, \tag{5}$$

where $\langle f(\mathbf{r},t)f(\mathbf{r}',t)\rangle=\delta(t-t')v(\mathbf{r}-\mathbf{r}')$. This describes a system of bosons of mass m interacting with potential

1063-651X/94/50(4)/2411(4)/$06.00

$$U = -\sum_{i<j}^{n} v(\mathbf{r}_i - \mathbf{r}_j) - \frac{n}{2} v(0) \tag{6}$$

evolving in imaginary time.

Equation (5) can equally well be interpreted in terms of a lattice model, for example the Hubbard model. The points $\mathbf{r}$ and ∇ are both discretized, which is the case studied numerically in the following sections. In this case the field ϕ is stored as a vector and evolved in time according to the above equation. The wave function at time t is

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_n) = \langle \phi(\mathbf{r}_1) \cdots \phi(\mathbf{r}_n) \rangle. \tag{7}$$

Note that the points $\mathbf{r}_1, \ldots, \mathbf{r}_n$ are kept fixed during the simulation and do not change as they would in the world line Monte Carlo method. It is the ϕ's that change their value.

To simulate n fermions, n fields $\phi_1, \phi_2, \ldots, \phi_n$ are evolved starting from different initial conditions but with the same $f(\mathbf{r},t)$. The wave function at time t in this case can be expressed in terms of the average of a Slater determinant:

$$\Psi(\mathbf{r}_1, \ldots, \mathbf{r}_n) = \left\langle \begin{vmatrix} \phi_1(\mathbf{r}_1) & \cdots & \phi_1(\mathbf{r}_n) \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \cdot & & \cdot \\ \phi_n(\mathbf{r}_1) & \cdots & \phi_n(\mathbf{r}_n) \end{vmatrix} \right\rangle. \tag{8}$$

Note that the potential in Eq. (6) is the *negative* of the correlation function. Therefore to simulate repulsive potentials requires an $f(\mathbf{r},t)$ that is complex [2]. This method is closely related to a simulation method for the density matrix [3,4] and also to the projector Monte Carlo method [5].

Obtaining the ground state energy for quantum systems such as the ones described above requires determining the asymptotic exponential decay of the wave function. In principal this can be done by averaging over many realizations of f; however, there are two reasons why this is impractical.

The first is a problem of importance sampling. As an example, consider determining the ground state energy $E(n)$ of n bosons. This is obtained by calculating $\langle \phi^n(\mathbf{r},t) \rangle$ and fitting this to an exponential $\exp[E(n)t]$ for long times. The problem is that $E(1) \neq E(n)/n$. As a result the typical behavior of $\phi^n(\mathbf{r},t)$ and the average behavior are very different. What dominates the average are some very rare realizations of f. Another way of describing such behavior is in terms of the notion of intermittency [6]. In order to get the correct average, one must use importance sampling. There are two ways to do this. One is to use a Monte Carlo method, keeping track of all time paths [3]. The second is to use a replication algorithm [7,8]. In order to keep track of angles, it is much more convenient to use the second method.

A large number of copies of systems described by Eq. (5) are run in parallel. The number of configurations that are replicated is proportional to their weight. For example, if we wish to measure $\langle |\phi|^{n+1} \rangle / \langle |\phi|^n \rangle$ at some site $\mathbf{r}$, we should choose a weight proportional to $|\phi|^n$. The proportionality factor is chosen so that the number of copies of the system stays almost constant. This method is similar to others described in detail in Refs. [7,8].

The second problem is one encountered when simulating repulsive bosons and fermions. Here nonpositivity leads to the problem discussed in the Introduction, that an exponentially large number of runs must be taken to obtain adequate statistics. The phase method proposed here is designed to reduce the number of runs. This will be illustrated numerically for a system of repulsive bosons. One should note that although it is possible to choose a method for this problem where there are no minus signs [9,10], this is an instructive example. It will be seen that the method used generalizes to fermions.

III. OBTAINING THE PHASE

Consider n bosons on a one dimensional lattice with L sites interacting via an on-site repulsive interaction U. This has been the focus of recent investigation [11] using the world line Monte Carlo method [9,10], which is clearly better suited to this problem than the method here. We use the method above only to illustrate techniques to cope with negative signs. In obtaining ground state energies, any component of $|\Psi\rangle$ can be measured. As will be seen below, in order to make the phase method work it is important to make a judicious choice for this. We start by choosing to measure $\Psi(x_1, x_2, \ldots, x_n)$ where the x_i's are all different sites. Later we will see why using the same sites does not work as well. The prescription for computing Ψ is given by Eq. (7).

The replication method outlined above was the case where all the x_i's are the same. To generalize the discussion above, one just chooses a weight $w \equiv |\phi(x_1)\phi(x_2)\cdots\phi(x_n)|$. The ground state energy of $n+1$ particles can be easily computed by knowing

$$\frac{\langle \phi(x_1)\cdots\phi(x_{n+1}) \rangle}{\langle |\phi(x_1)\cdots\phi(x_n)| \rangle} = \langle e^{i\theta} | \phi(x_{n+1}) | \rangle_w = \int P(\Theta,t) e^{i\Theta} d\Theta, \tag{9}$$

where the last average is done with respect to the weight w, and θ is the total phase angle of $\phi(x_1)\cdots\phi(x_{n+1})$. $P(\Theta,t) \equiv \langle |\phi(x_{n+1})| \delta(\Theta - \theta(t)) \rangle_w$ is closely related to $P(\theta,t)$ described below Eq. (3).

To test this out numerically, the number of lattice sites was chosen to be 8, and we started out by using a small value for $U = 0.5$. An average of 4497 copies were replicated. From this data $P(\Theta,t)$ was obtained with good statistics. In order to obtain an improved estimate of the tail of this distribution, the replication algorithm was altered to weight large Θ, that is by letting $w \to w\exp(\text{const}\Theta)$. By changing the constant in the exponent, $P(\Theta,t)$ was probed for different regions of Θ. By combining these, one obtains a good estimate for $P(\Theta,t)$ over nine orders of magnitude. An example of $\ln P(\Theta,t)$ is shown in Fig. 1 for $n=8$. The solid line is a quadratic fit to the data. By taking the Fourier transform of this for different times, as prescribed by Eq. (9), the ground state energy can be obtained. This is shown in Fig. 2 by the solid triangles. The open squares show the exact results for comparison.

Without using $P(\Theta,t)$, one can compute the weighted average of $|\phi(x_{n+1})|\exp(i\Theta)$ directly. For comparison the average of this, over the same copies, is shown Fig. 3. It is

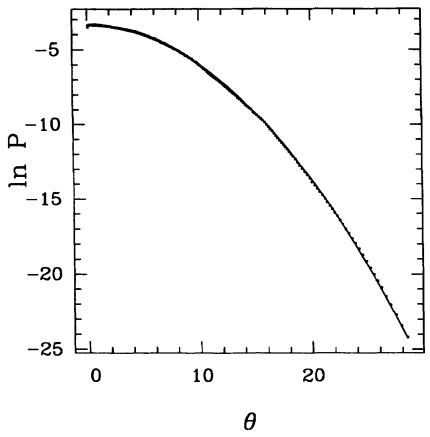

FIG. 1. The histogram $P(\theta,t)$ at $t=5$. The histogram was obtained by using the replication algorithm in conjunction with Eq. (5). The algorithm was run by giving favorable weights to large angles enabling good statistics for large values of the angle to be obtained. The solid line is a quadratic fit to the data.

clear that results obtained by direct averaging are far inferior to those obtained using $P(\Theta,t)$.

As can be seen in Fig. 1 the histogram can be obtained with high precision; therefore the statistical errors produced in obtaining the energy estimate are quite small. Therefore the difference between the exact and computed energies is greater than to be expected due to statistical error. This is due to the assumption of a Gaussian form for $P(\Theta,t)$. This will become much more apparent for the case of large U.

IV. IMPROVEMENTS TO THE METHOD

The method as it stands has difficulties for large U. The histogram of Θ has too large a variance. This can be understood by considering a variant of Eq. (1) in which there are two independent random processes with different f's,

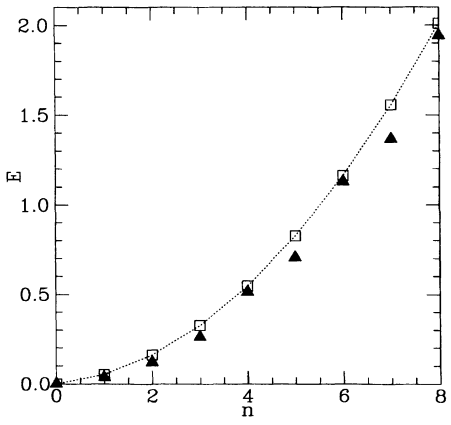

FIG. 2. The ground state energy E, as a function of number of particles n, for repulsive bosons with interaction energy of 0.5. The solid triangles are the results obtained using $P(\Theta,t)$. The open squares with the dashed line going through them are the exact results.

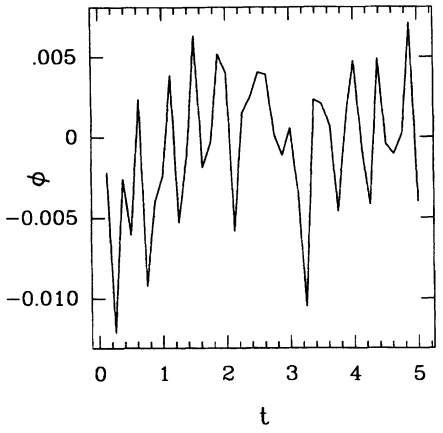

FIG. 3. The result of averaging the field over the same runs as was used to obtain Fig. 1. An exponential decay is not discernible and many more averages would be needed to see it. The real part is plotted here.

$\langle f_1(t)f_1(t')\rangle = v_1\delta(t-t')$, and $\langle f_2(t)f_2(t')\rangle = v_2\delta(t-t')$. The equations corresponding to each are $\dot{z}_i = if_i(t)z_i$, and therefore $z_i = e^{i\theta_i(t)}$, $i=1,2$. Suppose we are interested in numerically computing the sum

$$\langle A_1 z_1^n + A_2 z_2^n\rangle = A_1 e^{-n^2 v_1 t/2} + A_2 e^{-n^2 v_2 t/2}. \tag{10}$$

This average will be dominated for long times by the z_i with the smaller of the v_i's, say v_1. Under what circumstances will the numerical method of finding the histogram for the phase angle of the above sum give a sensible answer? Equation (10) can be thought of as the sum of two randomly rotating vectors with magnitudes $|A_i|$. From this observation, it is easy to see that, for long times, the variance of the phase angle will be the same as that for z_1^n if $|A_1|>|A_2|$. For $|A_1|<|A_2|$ a broader distribution will be obtained. Unless very careful fitting to a nonquadratic expression is done, an erroneously high energy will be obtained.

This example shows that very careful fitting is necessary if the wave function has too high an amplitude of excited states. One way to reduce this effect is to choose to measure Ψ on different sites as was done in the numerical example above. Another way to improve this method is to *filter* ϕ in time. It is easily shown that applying a linear filter to $\phi(t)$, $\gamma(t)\equiv\phi(t)\circ g(t)$ does not change the ground state energy if the filtering function $g(t)$ decays sufficiently rapidly. For example, one can evolve the equation

$$\dot{\gamma}(\mathbf{r},t) = -c\gamma(\mathbf{r},t)+\phi(\mathbf{r},t) \tag{11}$$

simultaneously with evolving Eq. (5) and measure γ^n instead of ϕ^n. This method works quite well. The energies for an eight site lattice for the case $U = 4$ are shown in Fig. 4. The energies are in good agreement with the exact results up to $n = 6$ after which it becomes apparent that the higher energy states start to dominate the answer. The method will presumably improve if in addition to filtering, different sites are used as in the example above. This filtering has similarities with the stationary phase Monte Carlo method [12] (SPMC), where a convolution of the integrand of the parti-

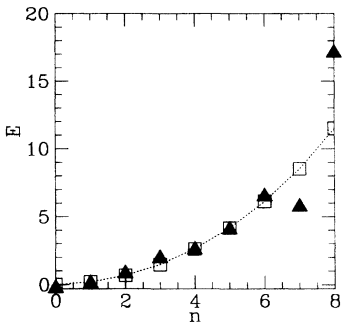

FIG. 4. The ground state energy E, as a function of number of particles n, for repulsive bosons with interaction energy of 4.0, obtained using linear filtering of the field (solid triangles), and the exact energy (open squares with the dashed line going through them).

tion function with a highly peaked function reduces the severity of the sign problem. However, the filtering above occurs only in the time domain, as opposed to SPMC where a convolution is performed over all variables of integration.

Lastly, the case of fermions has been considered. The determinant formulation of this problem can be used to obtain a continuous phase as a function of time. The energies for spinless fermions on three sites interacting with nearest neighbor repulsion U of 0.5, 2.0, and 8.0 are shown in Table I. Even for large U they agree well with the exact result [13].

TABLE I. The ground state energies for spinless fermions on three sites, as a function of the number of particles n, obtained using the phase method and the exact results.

	$U=0.5$		$U=2.0$		$U=8.0$	
n	E	exact	E	exact	E	exact
2	3.57	3.5	5.12	5	10.9	11
3	7.45	7.5	12.1	12	30.0	30

In conclusion, an alternative method for dealing with cancellations in quantum systems has been developed and tried out in some simple cases. Further theoretical development of the determination of $P(\Theta,t)$ could greatly improve fitting. Recent theoretical work [14] relating the phase used here to the Berry phase for smooth paths should be further explored in the context of this simulation method. It would also be interesting to consider more difficult problems such as the two dimensional Hubbard model with repulsive interactions using this method.

ACKNOWLEDGMENT

The author thanks Eliot Dresselhaus, Richard Scalettar, and Peter Young for useful discussions. This research was supported by the NSF under Grant No. DMR-9112767.

[1] S. Sorella, Int. J. Mod. Phys. B **5**, 937 (1991).

[2] Using complex fields is not necessary but useful in the fermion case as it promotes tunneling between configurations with different signs. See S.R. White, R.L. Sugar, and R.T. Scalettar, Phys. Rev. B **38**, 11 665 (1988).

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

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

[5] R. Blankenbecler and R.L. Sugar, Phys. Rev. D **27**, 1304 (1983).

[6] G. Paladin and A. Vulpiani, Phys. Rep. **156**, 147 (1987).

[7] D.M. Ceperley and M.H. Kalos, in *Monte-Carlo Methods in Statistical Mechanics*, edited by K. Binder (Springer Verlag, Berlin, 1979).

[8] D. Kung, R. Blankenbecler, and R.L. Sugar, Phys. Rev. B **32**, 3058 (1985).

[9] J.E. Hirsch, R.L. Sugar, D.J. Scalapino, and R. Blankenbecler, Phys. Rev. B **26**, 5033 (1982).

[10] M. Suzuki, K. Miyashita, and A. Kuroda, Prog. Theor. Phys. **58**, 1377 (1977).

[11] G.G. Batrouni, R.T Scalettar, and G.T. Zimanyi, Phys. Rev. Lett. **65**, 1765 (1990).

[12] J.D. Doll, D.L. Freeman, and M.J. Gillan, Chem. Phys. Lett. **143**, 277 (1988); J.D. Doll, T.L. Beck, and D.L. Freeman, J. Chem. Phys. **89**, 5753 (1988).

[13] In this case, the $nv(0)/2$ term in Eq. (6) was obtained directly from the input parameters and not obtained from measurements. This appears to give slightly better agreement with the exact results. Why the method works so well for large U in this fermionic example deserves further investigation.

[14] J.H. Samson, Phys. Rev. B **47**, 3408 (1993).