Guided Random Walks for Solving Hamiltonian
Lattice Gauge Theories*

S. A. Chin
Department of Physics
University of California
Los Angeles, CA 90024

J. W. Negele
Center for Theoretical Physics
Laboratory for Nuclear Science and Department of Physics
Massachusetts Institute of Technology
Cambridge, Massachusetts 02139

S. E. Koonin
Kellogg Radiation Laboratory
California Institute of Technology
Pasadena, CA 91125

* This work is supported in part by NSF-82-08439, DE-AC02-76ER03069,
NSF-PHY82-07337 and PHY82-15500

Submitted to Annals of Physics

November, 1983

CTP # 1127

Proposed running head: RANDOM WALKS FOR GAUGE THEORIES

S. A. Chin
Department of Physics
University of California
Los Angeles, CA 90024

LIST OF SYMBOLS

δ = del, l.c.
 ψ = psi, l.c.
 ϕ = phi, l.c.
 Δ = delta, u.c.
 ζ = zeta, l.c.
 $\mathcal{O}$ = script "oh"
 τ = tau, l.c.
 ρ = rho, l.c.
 ϵ = epsilon, l.c.
 δ = delta, l.c.
 ω = omega, l.c.
 Ω = omega, u.c.
 ∇ = del, u.c.
 λ = lambda, l.c.

CERN LIBRARIES, GENEVA

CM-P00066761

Motivated by developments for many-particle quantum systems, a Monte Carlo method for solving Hamiltonian lattice gauge theories without fermions is presented in which a stochastic random walk is guided by a trial wavefunction. To the extent that a substantial portion of the local structure of the theory can be incorporated in the trial function, the method offers significant advantages relative to existing techniques. The method is applicable to the study of $SU(N)$ lattice gauge theories, and its utility is demonstrated by solving the compact $U(1)$ gauge theory in three spatial dimensions.

1. INTRODUCTION

The Monte Carlo method provides a powerful tool for investigating the properties of quantum systems with many degrees of freedom. In this method, an observable of interest is expressed in terms of an integral containing a weight function which may be the ground state wavefunction, the partition function, or a discrete approximation to a functional integral. A random walk is then defined to generate an ensemble of statistically independent elements distributed according to this weight function, and the observable is evaluated stochastically subject only to controllable sampling errors. In its primitive form, the method has no way of utilizing one's physical understanding of the solution, and the numerical effort is expended indiscriminately in reproducing those aspects of the solution which may already be qualitatively understood as well as those which cannot be calculated any other way.

A crucial development in treating ground state properties of quantum many-body systems was the use of an approximate trial wave function by Kalos and collaborators to guide the random walk for the exact wavefunction [1, 2]. (Although this technique is called "importance sampling" in the many-body and "classical" Monte Carlo literature, we shall avoid this terminology since it has come to have a different meaning in the field theory literature.) This technique is appealing physically in that it incorporates known physics into the trial function and only requires the stochastic process to deal with that portion of the problem which is not understood. It has also proven to be of great practical utility in a broad range of application, including liquid helium [3-5], molecules [6-8], the electron gas [10], liquid and solid hydrogen [11], and nuclear systems [12-15, 24]. In all of these applications, the

statistical variance of calculations with a trial function is dramatically reduced relative to that of those without a trial function, usually to such an extent that the calculations would not have been feasible without the trial function. Of course, final results are, in principle, independent of the bias introduced by the trial function.

In this paper, we develop an analogous technique for the study of Hamiltonian field theories on a discrete spatial lattice. (This objective was first considered by D. Stump [9] using a variant technique.) The Hamiltonian form of a field theory on a spatial lattice has several advantages relative to the Lagrangian form on a space-time lattice. It offers the physical insight of dealing with a wavefunction and the parameters of a physically motivated trial function can be easily determined variationally. The random walk in field configurations will involve only a three-dimensional rather than a four-dimensional matrix, enabling the treatment of larger spatial lattices. The continuum time limit is easily evaluated, whereas the corresponding extrapolation in the Lagrangian form would be computationally impractical. Finally, ground state expectation values of propagators and other operators with finite-time extent may also be evaluated to obtain masses of low-lying excited states.

As a first application of this method to lattice gauge theory, we study the case of compact $U(1)$ [16-19] in three spatial dimensions. This example is sufficiently simple that the essence of the method can be displayed with a minimum of technical complications. Nevertheless, the presence of a phase transition [20-22] provides nontrivial structure and enables us to investigate the efficiency of a simple trial function in two differing phases. In addition to the ground state energy, which is a relatively insensitive test of the wavefunction, we also calculate the plaquette energy (the derivative of the

ground state energy with respect to the coupling constant) and also spatial Wilson loops of various size.

The paper is organized as follows. Section 2 is a pedagogical exposition of the Monte Carlo method for the case of a single particle Hamiltonian $H = p^2/2m + V(x)$, with obvious generalization to the case of N -degrees of freedom. An effort is made to motivate our particular stochastic procedure and to compare and contrast our treatment with related approaches. In Section 3, we review some relevant facets of the $U(1)$ lattice gauge theory and realize the lattice Hamiltonian in a form corresponding to a system of nonrelativistic, interacting bosons on a unit circle. The choice of the trial function is next discussed in Section 4 where we show that a product of functions of a single plaquette variable embodies substantial aspect of the physics in both the strong and the weak coupling regime. Numerical results for the ground state energy, the plaquette energy, and the size-dependence of spatial Wilson loops are presented in Section 5. Finally, conclusions and future applications to $SU(N)$ lattice gauge theories are discussed in Section 6.

2. MONTE CARLO METHOD

The essential elements of the stochastic method to be used for Hamiltonian lattice gauge theory can be demonstrated for the simple case of a single particle governed by the Hamiltonian $H = p^2/2m + V(x)$. The method will be described first intuitively as a diffusion process in the presence of a source and sink, and then algorithms for unguided and guided random walks will be derived from appropriate functional integrals

A. Diffusion Equation for the Ground State

One way of obtaining the ground state of H is to evolve the imaginary-time Schrödinger equation

$$-\frac{\partial}{\partial t} \psi(x,t) = \left[-\frac{1}{2m} \frac{\partial^2}{\partial x^2} + V(x) - E \right] \psi(x,t) \quad (2.1)$$

For any initial wavefunction $\psi(x)$ not orthogonal to the ground state $\psi_0(x)$, evolution exponentially damps all excited state components in $\psi(x)$ and, in the large t limit, $\psi(x,t)$ becomes proportional to $\psi_0(x)$. The norm of $\psi(x,t)$ at large time grows, decays, or remains constant depending upon whether E is greater than, lesser than, or equal to the ground state energy E_0 .

Since $\psi_0(x)$ and $\phi(x)$ can always be chosen to be real and positive, we can regard $\psi(x,t)$ at all times as a probability distribution. (Note that it is ψ , not ψ^2 , which is treated as a probability.) In the case of no potential and $E = 0$, eq. (2.1) is a simple diffusion equation and can be solved by a Gaussian random walk. As shown by Einstein [23], if an ensemble of points $\{x_i\}$ is initially chosen with probability distribution $\psi(x,t)$, then their distribution

at a later time $\psi(x,t+\Delta t)$ can be obtained by evolving each point in the ensemble according to

$$x_i \rightarrow x_i + \xi_i \quad (2.2)$$

where ξ_i is a Gaussian distributed random variable with zero mean and variance

$$\langle \xi_i^2 \rangle = \Delta t/m$$

For any finite Δt , this Gaussian random walk yields the exact distribution in the limit of large ensemble size.

In the special case with no kinetic term, eq. (2.1) becomes a simple rate equation with a source-sink term $E-V(x)$. Evolving this equation for a time Δt yields

$$\psi(x,t+\Delta t) = e^{\Delta t[E-V(x)]} \psi(x,t) \quad (2.3)$$

so that the wavefunction grows or decays at each point x by a factor $w(x) = e^{\Delta t[E-V(x)]}$. In the classically allowed or forbidden regions, $E-V(x)$ is respectively a (positive) source or a (negative) sink term which increases or decreases the density of points in the ensemble.

A simple method for changing the density of points according to $w(x_i)$ is to replicate each x_i by a factor $w(x_i)$, treating its fractional part as a probability. For example, in the case $w(x_i) < 1$, the point x_i is replicated 1 time (that is, retained in the ensemble) with probability $w(x_i)$ and is replicated 0 times (that is, deleted) with probability $1-w(x_i)$. Similarly, for the case $1 < w(x_i) < 2$, the point x_i is replicated twice (that is, one additional element x_i is added to the ensemble) with probability $w(x_i)-1$ and

replicated once with probability $1 - \{w(x_i) - 1\}$. In general, denoting the greatest integer in w by $[w]$, x_i is replicated $1 + [w(x_i)]$ times with probability $w(x_i) - [w(x_i)]$ and $[w(x_i)]$ times probability $1 - w(x_i) + [w(x_i)]$. Selecting a random number ζ uniformly distributed in the interval $[0, 1]$, replication factors having these probabilities may be concisely written as [8]

$$N_i = [w(x_i) + \zeta]$$

In the limit of sampling a large number of ζ 's, the average number of replications of x_i will be $w(x_i)$ so that the evolution required by eq. (2.3) is produced stochastically.

For sufficiently small Δt , the entire eq. (2.1) may thus be evolved by a sequential process of Gaussian random walks followed by replications. (This assertion will be made precise in the next section.) Repeated application of this two-step process will therefore evolve the ensemble approximating the distribution $\psi(x, t)$ to large time. If one were to draw trajectories of the ensemble points in space-time, the result would be a superposition of tree-like graphs. At each time step, each ensemble point would move a distance $\sim \sqrt{\Delta t/m}$ in a random direction. In the classically allowed region, each branch of the trajectories would sprout new branches as a result of replication, whereas, in the classically forbidden region, branches would terminate due to deletions. The role of both the kinetic energy in spreading out the wavefunction and the potential energy in concentrating the wavefunction in the region of greatest attraction are clearly evident in this branched random walk.

To conclude this intuitive discussion of the Monte Carlo method, several remarks are appropriate. 1) The ground state energy E_0 may be determined from

the time dependence of the ensemble size at large time. Once the ensemble has evolved long enough to relax to the ground state, E_0 may be determined as that value of E in eq. (2.1) which keeps the population stable, or alternatively from the ratio of population sizes at $t+T$ and t

$$\frac{N(t+T)}{N(t)} = \exp(-T(E_0 - E))$$

The ground state energy determined in this way will henceforth be referred to as the normalization energy and denoted as E_N . 2) After the stochastic random walk has relaxed to the ground state, continued evolution yields statistically independent ensembles distributed according to $\psi_0(x)$. (The number of time steps required for statistical independence of an observable O may be determined from computing the decay time τ_0 of the autocorrelation function: $[\langle O(t+\tau)O(t) \rangle - \langle O(t) \rangle^2] / [\langle O(t) \rangle^2 - \langle O(t) \rangle^2]$.) Thus arbitrarily good statistics for an observable may be obtained with a finite ensemble size by continuing the evolution sufficiently long and accumulating the values of $\{x_i\}$ after each interval τ_0 . 3) Systematic errors associated with the finite time step Δt may be eliminated by repeating the evolution for a sequence of successively smaller step sizes and extrapolating to the $\Delta t = 0$ limit. 4) In principle, the large time equilibrated, ensemble distribution will display all of the symmetries of the potential. The replication factor $\exp(\Delta t[E - V(x_i)])$ clearly has the symmetries of the potential and the Gaussian random walk will eventually remove any non-symmetric components in the initial wavefunction. However, in cases where the potential is degenerate (or nearly so), with minima separated by high barriers, the ensemble may take an impractically long time to equilibrate.

B. Unguided Random Walks from Path Integrals

Monte Carlo methods for calculating ground state or thermal properties of many-body systems are essentially stochastic means of evaluating observables as multiple integrals, and variant methods correspond to alternative choices of the multiple integral representation and stochastic sampling technique. In this section we show how a particular set of these choices leads to the random walk algorithm described in the preceding section.

The first choice concerning observables is whether to calculate thermodynamic properties at finite temperature, in which case one must sample $\text{tr } e^{-\beta H}$, or whether to calculate ground state properties, in which case one samples the ground state wavefunction $|\psi_0\rangle$. Our objective in this work will be to calculate the ground state properties.

The next choice is which operator to use to filter $|\psi_0\rangle$ from an initial wavefunction $|\phi\rangle$. The two most common choices are the resolvent operator $(E-z)(H-z)^{-1}$, with fixed $z < E_0$, from which $|\psi_0\rangle$ can be obtained via

$$|\psi_0\rangle = \lim_{N \rightarrow \infty} \left(\frac{E-z}{H-z} \right)^N |\phi\rangle, \quad (2.4)$$

and the evolution operator $e^{-\beta(H-E)}$, for which

$$\begin{aligned} |\psi_0\rangle &= \lim_{\beta \rightarrow \infty} e^{-\beta(H-E)} |\phi\rangle \\ &= \lim_{\substack{\Delta t \rightarrow 0 \\ N\Delta t \rightarrow \infty}} e^{-N\Delta t(H-E)} |\phi\rangle. \end{aligned} \quad (2.5)$$

By inserting complete sets of coordinate eigenstates between operators in (2.4) and (2.5), distinct multiple integral representations for $|\psi_0\rangle$ are

obtained. In the first case, (2.4), corresponding to the Green's function Monte Carlo method [1-11], each new ensemble $\{x'\}$ necessitates the use of an additional Neumann-Ulam random walk to sample $\langle x' | (H-z)^{-1} | x \rangle$ stochastically. In the second case, (2.5), each new ensemble requires only the calculation of Gaussian steps and replication factors, but the entire walk must be repeated several times in order to extrapolate to the $\Delta t = 0$ limit. As a result, the numerical effort in both methods is roughly comparable. In this work, we have chosen the path integral form because of its computational simplicity, its physical appeal, and the ease with which observables can be expressed in terms of path integrals.

In principle, complete sets of states other than coordinate eigenstates, such as momentum eigenstates, boson or Grassman coherent states, and permanents or Slater determinants, can also be used. For our present purpose, the choice of coordinate eigenstates is most appropriate.

Finally, one must decide how to evaluate stochastically the resulting multiple integral. Consider the calculation of the overlap of the ground state $|\psi_0\rangle$, given by (2.5), with a state $|\chi\rangle$. For an infinitesimal time step Δt ,

$$\begin{aligned} \langle x_{n+1} | e^{-\Delta t(H-E)} | x_n \rangle &\approx \langle x_{n+1} | e^{-\Delta t p^2/2m} e^{-\Delta t[V(x)-E]} + O(\Delta t^2) | x_n \rangle \\ &= \int dp \langle x_{n+1} | e^{-\Delta t p^2/2m} | p \rangle \langle p | e^{-\Delta t[V(x)-E]} | x_n \rangle \\ &\approx \left(\frac{m}{2\pi\Delta t} \right)^{1/2} \exp \left[-\frac{m}{2\Delta t} (x_{n+1} - x_n)^2 \right] \exp \left[-\Delta t[V(x_n) - E] \right] \\ &\equiv P(x_{n+1}, x_n) W(x_n), \end{aligned} \quad (2.6)$$

so that

$$\begin{aligned}
\langle \zeta | e^{-N\Delta t(H-E)} | \phi \rangle &= \\
&= \int dx_N \dots dx_0 \langle \zeta | x_N \rangle \langle x_N | e^{-\Delta t(H-E)} | x_{N-1} \rangle \dots \langle x_2 | e^{-\Delta t(H-E)} | x_1 \rangle \langle x_1 | e^{-\Delta t(H-E)} | x_0 \rangle \langle x_0 | \phi \rangle \\
&= \int dx_N \dots dx_0 \zeta(x_N) P(x_N, x_{N-1}) w(x_{N-1}) \dots P(x_2, x_1) w(x_1) P(x_1, x_0) w(x_0) \phi(x_0). \quad (2.7)
\end{aligned}$$

Given a set of points $\{x^i\}$ distributed according to a probability distribution function $P(x)$, integrals of the form $\int W(x)P(x)dx$ can be evaluated by the Monte Carlo method as [26]

$$\int W(x) P(x) dx = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N W(x^i). \quad (2.8)$$

Hence, a simple stochastic method for evaluating the multiple integral (2.7) is as follows. First, an ensemble of points $\{x_0^i\}$ is selected according to the distribution $\phi(x_0)$, for example, using the Metropolis algorithm [25]. Next, corresponding to each x_0^i in the ensemble, a sequence of points $x_1^i, x_2^i, \dots, x_N^i$ is generated with each x_{s+1}^i successively chosen to be Gaussian distributed around x_s^i in accordance with the distribution $P(x_{s+1}, x_s^i)$. The resulting ensemble of sequences, $\{x_0^i, x_1^i, x_2^i, \dots, x_N^i\}$, with each member defining a particular path, is thus distributed with probability $\prod_{s=0}^{N-1} P(x_{s+1}, x_s^i) \phi(x_0^i)$. The path integral (2.7) can therefore be evaluated according to (2.8) as

$$\langle \zeta | e^{-N\Delta t(H-E)} | \phi \rangle = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \zeta(x_N^i) \prod_{s=0}^{N-1} w(x_s^i). \quad (2.9)$$

However, this procedure is in general inefficient in that the accumulating weight $\prod_{s=0}^{N-1} w(x_s^i)$ can vary dramatically from path to path resulting in an unnecessarily high variance as well as waste of labor in evaluating those paths with insignificant weights. The algorithm can be improved by generating each x_{s+1}^i successively according to the entire matrix element $P(x_{s+1}, x_s^i) w(x_s^i)$ and thus eliminating all weights $w(x_s)$ from (2.9). This then reproduces the branched random walk described in the preceding section whereby each point x_s^i is Gaussian random walked according to $P(x_{s+1}, x_s)$ and replicated stochastically by a factor $w(x_s^i)$. In this case, the final ensemble $\{x_N^i\}$ is distributed according to $\prod_{s=0}^{N-1} P(x_{s+1}, x_s^i) w(x_s^i) \phi(x_0)$ and (2.9) simplifies to

$$\langle \zeta | e^{-N\Delta t(H-E)} | \phi \rangle = \lim_{N \rightarrow \infty} \frac{1}{N} \sum_{i=1}^N \zeta(x_N^i).$$

Given this stochastic method for evaluating matrix elements of the evaluation operator, a number of observables of physical interest can be evaluated straightforwardly. In addition to the normalization energy, E_N , defined previously, the ground state energy can also be calculated directly from the wavefunction via

$$\begin{aligned}
E_0 &= \lim_{\Delta t \rightarrow 0} \frac{\langle H\phi | e^{-N\Delta t(H-E)} | \phi \rangle}{N\Delta t \langle \phi | e^{-N\Delta t(H-E)} | \phi \rangle} \\
&= \frac{\sum_i H\phi(x_N^i)}{\sum_i \phi(x_N^i)} \equiv E_T. \quad (2.10)
\end{aligned}$$

The value of the energy determined in this way is referred to as the trial energy E_T . If ϕ were the exact ground state wavefunction, E_T would be exact irrespective of statistical or systematic errors in the x_N^i 's. In practice, a

good trial function significantly reduces the statistical variance and systematic bias in computing E_T . Since E_N and E_T are independent ways of determining E_0 , agreement between these two results in actual calculations constitutes a nontrivial consistency test of the entire Monte Carlo algorithm.

A set of points $\{x\}$ distributed according to $\psi(x)$ does not allow one to evaluate integrals of the form $\int f(x)\psi^2(x)dx$ directly. Thus the evaluation of ground state expectation of operators other than the Hamiltonian requires one of two alternative approaches. When ϕ is a good approximation to ψ_0 such that the derivation $\delta = \psi_0 - \phi$ is small, a simple perturbative estimate [2] is most convenient. To second order in δ , $\psi_0^2 = 2\phi\psi_0 - \phi^2 = \phi^2 + 2\phi\delta = \phi\psi_0 + \phi\delta$, and therefore a convenient perturbative estimate is

$$\begin{aligned} \frac{\langle \psi_0 | O | \psi_0 \rangle}{\langle \psi_0 | \psi_0 \rangle} &= \frac{2\langle \phi | O | \psi_0 \rangle - \langle \phi | O | \phi \rangle}{\langle \psi_0 | \psi_0 \rangle} + O(\delta^2) \\ &= 2 \frac{\langle \phi | O | \psi_0 \rangle}{\langle \phi | \psi_0 \rangle} - \frac{\langle \phi | O | \phi \rangle}{\langle \phi | \phi \rangle} + O(\delta^2) \\ &= 2 \frac{\sum_i \phi(x_N^i) \int \phi(x_0^i) O\phi(x_N^i)}{\sum_i \phi(x_N^i)} - \frac{\sum_i \phi(x_0^i) \int \phi(x_0^i) O\phi(x_0^i)}{\sum_i \phi(x_0^i)} \end{aligned} \quad (2.11)$$

assuming that the operator O is hermitian. The two terms on the RHS of (2.11) are evaluated in the same manner as the trial energy, with x_0^i 's drawn from the initial population $\{x_0\}$.

When necessary, the exact expectation value of a local operator $\langle x^i | O | x \rangle = \delta(x^i - x)O(x)$ may be calculated by evaluating the expression

$$\langle O \rangle = \lim_{\Delta t \rightarrow 0} \frac{\langle \phi | e^{-N\Delta t(H-E)} O e^{-N\Delta t(H-E)} | \phi \rangle}{\langle \phi | e^{-2N\Delta t(H-E)} | \phi \rangle} \quad (2.12)$$

In this case, at the N^{th} step, $O(x_N^i)$ is evaluated for each element x_N^i in the

ensemble $\{x_N\}$. For the remaining N steps of the random walk, a record is kept of the element of $\{x_N^i\}$ from which each subsequent ensemble member has evolved. At the end of the $2N^{\text{th}}$ step, $\langle O \rangle$ is evaluated by forming the weighted sum

$$\langle O \rangle = \frac{\sum_i \phi(x_{2N}^i) O(x_N^i)}{\sum_i \phi(x_{2N}^i)} \quad (2.13)$$

where the notation $i(j)$ indicates that i is the progenitor of j . Propagators and operators with finite time extent may also be evaluated analogously. For example, the low lying spectrum can be investigated in the usual way by computing the expectation value of operators at two different times and extracting the leading exponential dependence:

$$\begin{aligned} \langle O(T)O(0) \rangle &= \sum_n \exp[-(E_n - E_0)T] |\langle n | O | \psi \rangle|^2 \\ &= \frac{\langle \phi | e^{-N\Delta t(H-E)} O e^{-N\Delta t(H-E)} O e^{-N\Delta t(H-E)} | \phi \rangle}{\langle \phi | e^{-(2N+M)\Delta t(H-E)} | \phi \rangle} \end{aligned} \quad (2.14)$$

where $T = M\Delta t$ and N is again large. Descendants of each ensemble members of $\{x_N\}$ and $\{x_{N+M}\}$ must be followed to the final $2N+M^{\text{th}}$ step and weighted with $O(x_N^i)O(x_{N+M}^j)$. Since implementation of these exact operator expectation values is cumbersome, we have used the simpler perturbative estimates in this initial calculation. However, it is important to emphasize that these expectation values are within the purview of the present theory, and that the ultimate objective of developing the method presented here includes the computation of such observables.

C. Guided Random Walks

The Monte Carlo method described in the preceding section can be improved significantly by guiding the random walk using an approximate trial function $\phi(x)$ which contains as much of the essential physics as is understood at the outset. Instead of evolving $\psi(x,t)$ directly, one evolves the product wavefunction $\phi(x)\psi(x,t) \equiv \rho(x,t)$, which, by virtue of (2.1), satisfies the equation

$$-\frac{\partial \rho}{\partial t} = -\frac{1}{2m} \frac{\partial^2 \rho}{\partial x^2} + \frac{\partial}{\partial x} \left[\frac{1}{m} \phi^{-1} \frac{\partial \phi}{\partial x} \rho \right] + [\phi^{-1} H \phi - E] \rho. \quad (2.15)$$

Note that to the extent to which ϕ is a good approximation to ψ_0 , one is now sampling the probability distribution ϕ^2 instead of the wavefunction ψ_0 . In contrast to the unguided case, in which the replication factor $e^{\Delta t[E-V(x)]}$ gives rise to large fluctuation in the population and corresponding large variance in observables, the replication factor in the guided case is

$$w(x) = e^{\Delta t[E-\phi^{-1}(x)H\phi(x)]} \quad (2.16)$$

and deviates from unity only to the extent that $H\phi(x)$ differs from $E\phi(x)$. A reasonable trial function will therefore significantly reduce the amount of population fluctuation. The remaining terms of (2.15) describe diffusion in an external drift velocity field $v(x) = \frac{1}{m} \phi^{-1}(x) \frac{\partial \phi(x)}{\partial x}$, and one would intuitively expect them to be simulated, for small Δt , by the extension of

$$x_i \rightarrow x_i + \xi_i + \Delta t \frac{1}{m} \phi^{-1}(x) \frac{\partial \phi(x)}{\partial x}. \quad (2.17)$$

Thus the trial function guides the random walks away from region where ϕ is small toward regions where it is large. In the limit in which ϕ is exact, all of the evolution is contained in the Gaussian diffusion and the replication factor $w(x)$ is an irrelevant constant.

The algorithm for evolving an ensemble $\{x_s\}$ which samples ρ can be obtained by inserting the resolution of unity $1 = \int dx \frac{1}{\phi(x)} |x\rangle \langle x|$ at each infinitesimal time step in the imaginary-time evolution operator [14]

$$\rho(x,t) = \langle x | \phi \psi(t) \rangle = \langle x | \phi e^{-\Delta t(H-E)} \dots e^{-\Delta t(H-E)} | \phi \rangle \quad (2.5)$$

$$= \int dx_N \dots dx_0 \langle x | \phi e^{-\Delta t(H-E)} \frac{1}{\phi} | x_N \rangle \langle x_N | \phi e^{-\Delta t(H-E)} \frac{1}{\phi} | x_{N-1} \rangle \dots$$

$$\dots \langle x_1 | \phi e^{-\Delta t(H-E)} \frac{1}{\phi} | x_0 \rangle \langle x_0 | \phi \rangle. \quad (2.18)$$

The initial ensemble $\{x_0\}$ must now be distributed according to $\langle x_0 | \phi \rangle = \phi^2(x_0)$. The matrix element $\langle x_{N+1} | \phi e^{-\Delta t(H-E)} \frac{1}{\phi} | x_N \rangle$ needs to be evaluated with some care to avoid omission of terms of order Δt . Expanding the exponential in powers of Δt , commuting all momentum operators to the left of all position operators in each term, inserting a complete set of momentum eigenstates, and reexponentiating, gives

$$\begin{aligned}
\langle x_{n+1} | \phi e^{-\Delta t(H-E)} \frac{1}{\phi} | x_n \rangle &= \langle x_{n+1} | \left\{ 1 - \Delta t \left[\frac{1}{2m} \left(p^2 + 2ip \frac{\phi'}{\phi} \right) + \left(-\frac{1}{2m} \frac{\phi''}{\phi} + V - E \right) \right] \right. \\
&\quad \left. + \frac{1}{2} \Delta t^2 \left[\frac{1}{4m^2} \left(p^2 p^2 + 4ip^2 p \frac{\phi'}{\phi} \right) + \frac{p^2}{m} \left(-\frac{1}{2m} \frac{\phi''}{\phi} + V - E \right) - \frac{p^2}{m^2} \frac{\phi'''}{\phi} + O(p) + O(1) \right] + \dots \right\} | x_n \rangle \\
&= \int \frac{dp}{2\pi} \exp \left\{ -\frac{\Delta t}{2m} \left[1 + \frac{\Delta t}{m} \frac{d^2}{dx^2} \ln \phi(x_n) \right] p^2 + i \left[x_{n+1} - x_n - \frac{\Delta t}{m} \frac{\phi'(x_n)}{\phi(x_n)} + O(\Delta t^2) \right] p \right. \\
&\quad \left. - \Delta t \left[\phi^{-1}(x_n) H \phi(x_n) - E + O(\Delta t) \right] \right\} \\
&= \left[\frac{m}{2\pi \Delta t \phi(x_n)} \right]^{1/2} \exp \left\{ -\frac{m}{2\Delta t \phi(x_n)} \left[x_{n+1} - x_n - \frac{\Delta t}{m} \frac{\phi'(x_n)}{\phi(x_n)} \right]^2 \right. \\
&\quad \left. - \Delta t \left[\phi^{-1}(x_n) H \phi(x_n) - E \right] \right\} + O(\Delta t^2) \quad (2.19)
\end{aligned}$$

where $\tilde{\Delta t}(x) = \left[1 + \frac{\Delta t}{m} \frac{d^2}{dx^2} \ln \phi(x) \right]$.

In counting powers of Δt in the final result, it is important to recognize that the Gaussian integral over p implies that $p^2 \sim 1/\Delta t$ so that the term $\frac{\Delta t^2}{2m^2} \frac{d^2}{dx^2} \ln \phi(x)$ p^2 is really of order Δt and must be kept. In contrast, terms $\Delta t^2 p$ and Δt^2 may be legitimately neglected. Equation (2.19) has essentially the form expected from the drift and source terms in eq. (2.15).

The only unexpected result from the path integral derivation is that the random walk step size $\tilde{\Delta t}(x)$ in principle depends upon the value of x . In the special case of a Gaussian trial function, $\tilde{\Delta t}(x)$ reduces to an irrelevant change of time scale and in most practical applications it has negligible effect on the final result.

With an ensemble of points $\{x_n^i\}$ distributed according to $\rho(x)$, the calculation of observables reduces to a simple average, rather than the ratio of averages. For example, the trial energy now reads

$$E_T = \frac{\langle H \phi | \phi^{-1} \rho \rangle}{\langle \phi | \phi^{-1} \rho \rangle} = \frac{1}{N} \sum_{i=1}^N \phi^{-1}(x_N^i) H \phi(x_N^i) \quad (2.20)$$

As an estimator for E_0 , the above form for E_T completely eliminates the possible bias from dividing one fluctuating quantity by another. By making the replication factor close to unity, the use of a trial function not only reduces the observable variance, but it also lessens the bias from constantly adjusting E to stabilize the ensemble size and allows the calculation to be done with a smaller population. In practice, significant variance reduction can be achieved with any reasonable trial function and the method of guided random walks can be an efficient means of calculating the ground state energy.

For the perturbative estimate of the expectation value of a local operator O , (2.11) is replaced by

$$\langle O \rangle = 2 \frac{1}{N} \sum_{i=1}^N O(x_N^i) + \frac{1}{K} \sum_{i=1}^K O(x_0^i) \quad (2.21)$$

In contrast to the trial energy, the variance in $\langle O \rangle$ does not vanish as ϕ approaches ψ_0 and the statistical accuracy of a general operator is usually lower than for the energy. Nevertheless, reliable results can be obtained by finding consistent extrapolations given by (2.21) with different initial trial functions.

This concludes our introduction to the Monte Carlo method. Generalization to the case of N particles without the complication of antisymmetry is obvious: one simply replaces all the x 's by N -dimensional vectors $\vec{x}$ and the random walks take place in N -dimensional space. For example, the Gaussian evolution of the N -dimensional vector $\vec{x}_1$ is

$$\vec{x}_1 + \vec{x}_1 + \vec{x}_1 + \Delta t \frac{1}{m} \vec{\nabla} \phi(\vec{x}_1) / \phi(\vec{x}_1) \quad (2.22)$$

with replication factor

$$w(\vec{x}_1) = \exp \left(\Delta t [E - H \phi(\vec{x}_1) / \phi(\vec{x}_1)] \right) \quad (2.23)$$

3. APPLICATION TO U(1) LATTICE GAUGE THEORY

A. Lattice Hamiltonian

In a three dimensional simple cubic lattice with lattice constant a , the link connecting site $\vec{x} = (x_1, x_2, x_3)$ with site $\vec{x} + a\hat{e}_j$ can be uniquely identified by the site $\vec{x}$ and the direction j . A dynamical variable A defined on each link of the lattice may thus be labeled $A_j(\vec{x})$. In the U(1) lattice gauge theory, each directed link is associated with an element of the U(1) group defined by $A_j(\vec{x})$ via

$$U_{\pm j}(\vec{x}) = \exp[\pm i a g A_j(\vec{x})] \quad (3.1)$$

The plus or minus sign in eq. (3.1) specifies the direction of the given link; for example, $U_{-j}(\vec{x})$ is associated with the link $(\vec{x}, j)$ pointing in the minus j -direction. Since the group manifold of U(1) is a unit circle, (3.1) implies that $A_i(\vec{x})$ are angular variables with range $[-\pi/ga, \pi/ga]$.

The compact U(1) lattice Hamiltonian in the $A_0 = 0$ gauge is given by [17-19]

$$H = a^3 \sum_{i, \vec{x}} \left\{ \frac{1}{2} E_i^2(\vec{x}) + \frac{1}{2a^4} [1 - \cos(ga^2 B_i(\vec{x}))] \right\} \quad (3.2)$$

where $E_i(\vec{x})$ is the momentum conjugate to $A_i(\vec{x})$. The cosine term results from summing two oppositely directed, closed links around an elementary square or plaquette, each corresponding to a product of four U's. The resultant argument $B_i(\vec{x})$ is the lattice curl of $A_i(\vec{x})$,

$$B_i(\vec{x}) = \epsilon_{ijk} \nabla_j A_k(\vec{x}) \equiv \epsilon_{ijk} \frac{1}{a} [A_k(\vec{x} + a\hat{e}_j) - A_k(\vec{x})] \quad (3.3)$$

and may be associated with the plaquette at $\vec{x}$ spanned by the other two positive direction links perpendicular to the i -direction. The theory is quantized by imposing the canonical commutation relation:

$$[A_i(\vec{x}), E_i(\vec{x}')] = -i \frac{1}{a^3} \delta_{ii'} \delta_{\vec{x}\vec{x}'}, \quad (3.4)$$

In the continuum or weak coupling limit, $ga^2 \rightarrow 0$, (3.2) reduces to the conventional QED Hamiltonian without fermions.

By scaling out the lattice spacing a via $\vec{x} \rightarrow \vec{x}a$, $A \rightarrow A/ga$, $B \rightarrow B/ga^2$, $E \rightarrow gE/a^2$, $H \rightarrow g^2H/a$, we obtain a dimensionless Hamiltonian in the form

$$H = \sum_{i,\vec{n}} \frac{1}{2} E_i^2(\vec{n}) + \lambda \sum_{i,\vec{n}} [1 - \cos B_i(\vec{n})] \quad (3.5)$$

with

$$B_i(\vec{n}) = \epsilon_{ijk} [A_k(\vec{n} + \hat{e}_j) - A_k(\vec{n})], \quad (3.6)$$

$\vec{n} = a$ triplet of integers, $\lambda \equiv 1/g^4$, and $-\pi < A_i(\vec{n}) < \pi$. The commutation relation (3.4) can be realized in the space of square integrable functions of the link variable $A_i(\vec{n})$ via the conventional identification,

$$E_i(\vec{n}) = -i \frac{\partial}{\partial A_i(\vec{n})}.$$

In this case, one can regard the lattice Hamiltonian (3.5) as describing a system of nonrelativistic, unit mass bosons, each labeled by $(\vec{n}, i)$, residing on the unit circle with positions $\{A_i(\vec{n})\}$ and interacting with each other through a sum of four-body potentials $V(A_i(\vec{n})) = \lambda \sum_{i,\vec{n}} [1 - \cos B_i(\vec{n})]$. The underlying lattice structure then serves merely to identify participants of

each four-body interaction. For a periodic lattice with L^3 distinct sites, there are $N = 3L^3$ links. Hence, for a given lattice size L , (3.5) is equivalent to a N -boson problem. From this perspective, any method that can solve a quantum many-particle problem can solve the Hamiltonian gauge theory defined by (3.5) to (3.7). In particular, we can solve it by the method of guided random walks as described in the last section.

For an angular variable A with conjugate angular momentum E , the matrix element of the kinetic evolution operator reads

$$\begin{aligned} \langle A' | \exp[-\Delta t \frac{1}{2} E^2] | A \rangle &= \int_m \langle A' | \exp[-\Delta t \frac{1}{2} E^2] | m \rangle \langle m | A \rangle \\ &= \int_m \frac{1}{2\pi} \exp[i m(A' - A)] \exp[-\Delta t \frac{1}{2} m^2] \\ &= \int_m \frac{1}{\sqrt{2\pi\Delta t}} \exp[-\frac{1}{2\Delta t} (A' - A + 2\pi m)^2] \\ &\xrightarrow{\Delta t \rightarrow 0} \frac{1}{\sqrt{2\pi\Delta t}} \exp[-\frac{1}{2\Delta t} (A' - A)^2], \end{aligned} \quad (3.8)$$

so that in the limit of Δt goes to zero, A' is Gaussian distributed on the unit circle around the point A . Thus the details of the Monte Carlo algorithm are unaffected by the circular geometry of the $U(1)$ lattice gauge theory and (3.5) can be solved stochastically using (2.22) and (2.23). This is as one would have expected physically: in the limit of small Δt , the local process of diffusion should not be affected by the global structure of the manifold.

B. Gauge Invariance

Since the Hamiltonian (3.5) depends on $A_i(\vec{n})$ only through its lattice curl $B_i(\vec{n})$, it is independent of the lattice divergence of $A_i(\vec{n})$:

$$\vec{\nabla} \cdot \vec{A}(\vec{n}) \equiv \sum_i [A_i(\vec{n}) - A_i(\vec{n} - \vec{e}_i)]$$

Thus any transformation of $A_i(\vec{n})$, whose sole effect is to alter the value of $\vec{\nabla} \cdot \vec{A}(\vec{n})$, will leave the Hamiltonian invariant. This particular symmetry of H is gauge invariance. The required gauge transformations are obviously generated by the operator [18,19]

$$Q[\omega(\vec{n})] = \exp \left\{ i \sum_{\vec{n}} \omega(\vec{n}) \vec{\nabla} \cdot \vec{E}(\vec{n}) \right\} \quad (3.9)$$

with generators

$$\vec{\nabla} \cdot \vec{E}(\vec{n}) \equiv \sum_i [E_i(\vec{n}) - E_i(\vec{n} - \vec{e}_i)] \quad (3.10)$$

The effect of these transformations is to shift the three link variables connected to site $\vec{n}$ in the positive directions by $\omega(\vec{n})$ and the three link variables connected to site $\vec{n}$ in the negative directions by $-\omega(\vec{n})$. The net result is to shift each link variable by a discrete gradient: $A_i(\vec{n}) + A_i(\vec{n}) + \omega(\vec{n}) - \omega(\vec{n} + \vec{e}_i)$, which is the natural lattice analog of the continuum gauge transformation.

A function of the link variables, $\phi(A_i(\vec{n}))$, is gauge invariant if $Q\phi = \phi$. In this case, ϕ is annihilated by all the generators of (3.10):

$$\vec{\nabla} \cdot \vec{E}(\vec{n})\phi = 0 \quad (3.11)$$

Consequently, only gauge invariant eigenstates of H, which obey Gauss' Law (3.11), are physical. By applying Q to the operator equation (2.5) or (2.18),

which defines the Monte Carlo algorithm, and noting that $\vec{\nabla} \cdot \vec{E}(\vec{n})$ commutes separately with the kinetic and potential energy term of H, one can readily show that the resultant wavefunction is gauge invariant if the trial function is gauge invariant. In the unguided case, this is not an essential requirement. Since, as shown in Section 2.A, the stochastic solution will possess all the symmetries of the potential irrespective of the trial wave function. In the guided case, however, since ϕ occurs at every time step, the gauge invariance of ϕ is crucial.

By virtue of (3.11), a gauge invariant function $\phi(A_i(\vec{n}))$ is a function of only $2L^3$ independent variables. Integrating ϕ over the full set of $3L^3$ link variables will therefore yield a multiplicative gauge volume factor $V_G = (2\pi)^{L^3}$. In the continuum theory, the corresponding factor is divergent and must be eliminated by gauge fixing. For compact lattice gauge theory, since V_G is finite, this is unnecessary. Such a factor has no effect on evaluating expectation values of normalized wavefunctions and is never explicitly required or evaluated by the Monte Carlo algorithm.

4. THE TRIAL FUNCTION

In order to devise a reasonable trial function to guide the random walk, one must know as much as possible about the ground state of the lattice Hamiltonian (3.5). For this purpose, it is useful to examine two limiting cases where ψ_0 is known exactly.

In the strong coupling limit $g \rightarrow \infty$, and hence $\lambda \rightarrow 0$, perturbation theory in powers of λ gives

$$\psi_0 = 1 + \frac{1}{2} \lambda \sum_p \cos(B_p) + O(\lambda^2) \quad (4.1)$$

$$\epsilon_0 = \frac{E_0}{N} = \lambda - \frac{1}{4} \lambda^2 + \frac{3}{1280} \lambda^4 + O(\lambda^6) \quad (4.2)$$

where for convenience we have introduced the generic label "p" to denote the plaquette sum. This strong coupling result has the following simple physical interpretation [16,18]. When $\lambda = 0$, the unperturbed ground state wavefunction may be taken to be unity with zero energy. When λ is small, the potential term in (3.5) (ignoring the overall constant $N\lambda$) creates oriented single plaquette states with wavefunction $\exp[iB_p]$ and unperturbed energy $E_p = 2$.

The sum over all possible single plaquette excitations, with the appropriate energy denominator, changes the unperturbed ground state as indicated in (4.1). The shift in the ground state energy, with only single plaquette intermediate states, is then given by (4.2). Thus for λ small, the exact ground state is dominated by uncorrelated, single plaquette excitations, and we may rewrite (4.1) as a product of single plaquette functions:

$$\psi_0 = \prod_p \exp \left\{ -\frac{1}{2} \lambda [1 - \cos B_p] \right\} + O(\lambda^2) \quad (4.3)$$

with each plaquette function now normalized to one when its argument is zero.

In the weak coupling limit $g \rightarrow 0$, and hence $\lambda \rightarrow \infty$, the lattice Hamiltonian (3.5) can be rewritten in terms of the scaled variables: $\tilde{A} = \sqrt{\lambda} A$, $\tilde{B} = \sqrt{\lambda} B$, and $\tilde{E} = E/\sqrt{\lambda}$ to obtain the expansion

$$H = \lambda \left\{ -\frac{1}{2} \sum_p \tilde{E}_p^2 + \frac{1}{2} \sum_p \tilde{B}_p^2 - \frac{1}{4l} \sum_p \frac{1}{\lambda^2} \tilde{B}_p^4 + \dots \right\} \quad (4.4)$$

Retaining only the harmonic terms yields a set of independent harmonic oscillators with ground state wavefunction

$$\psi_0 = \exp \left\{ -\frac{1}{2} \frac{1}{\sqrt{\lambda}} \sum_{\vec{n}, \vec{n}'} \tilde{B}_1(\vec{n}) \tilde{B}_1(\vec{n}') \right\} \quad (4.5)$$

and energy

$$\epsilon_0 = C_0 \sqrt{\lambda} - \frac{1}{8} C_0^2 + O(1/\sqrt{\lambda}) \quad (4.6)$$

where

$$\Gamma(\vec{n}, \vec{n}') = \frac{1}{L^3} \sum_{\vec{k}} \left\{ 4 \sum_i \sin^2(K_i/2) \right\}^{-1/2} e^{i\vec{k} \cdot (\vec{n} - \vec{n}')} \quad (4.7)$$

$$C_0 = \frac{1}{3l^3} \sum_{\vec{k}} \left\{ 4 \sum_i \sin^2(K_i/2) \right\}^{1/2} \quad (4.8)$$

and $\vec{k} = 2\pi(k_1, k_2, k_3)/L$ with $k_i = 0, 1, 2, \dots, L-1$. The constant term in (4.6) as well as the λ^4 term in eq. (4.2) is taken from ref. 9. Also, the wavefunction (2.5) is not strictly correct; since the Hamiltonian (3.5) is periodic in B_p , so must be ψ_0 . Henceforth, we will tacitly assume that all such expressions are made periodic by summing over all integers n in the shifted plaquette variable $B_p + n2\pi$. The constant C_0 , depending on the lattice

size, has the values 0.7479, 0.7934, 0.7958, and 0.7959 for the values $L = 2, 4, 10$, and 50 , respectively. For the same lattice sizes, $\Gamma(0,0,0) = 0.3562$, 0.4281 , 0.4509 , 0.4552 ; $\Gamma(1,0,0) = -0.0178$, 0.0314 , 0.0529 , 0.0572 ; and $\Gamma(2,0,0) = 0.3562$, -0.0024 , 0.0110 , 0.0152 . Thus although in this limit the wavefunction (4.5) indicates that plaquettes are correlated, the correlations, as measured by the ratios $\Gamma(\vec{n})/\Gamma(\vec{0})$, are not very strong. For example, the nearest-neighbor ratio, $\Gamma(1,0,0)/\Gamma(0,0,0)$, for all lattice sizes, does not exceed 13% and the next nearest-neighbor ratio, $\Gamma(1,1,0)/\Gamma(0,0,0)$ does not exceed 6%. Therefore it is useful to guide the Monte Carlo calculation with a trial function having only the dominant diagonal terms, which is again a product wavefunction.

Thus, in this work we have chosen to approximate the ground state by a trial function of the product form

$$\phi = \prod_p f(\mathbf{B}_p) \quad (4.9)$$

for all values of λ . The requirements of gauge invariance and periodicity are easily satisfied by demanding $f(\mathbf{B})$ to be a periodic function of the plaquette variable. From the above discussion of the strong and weak coupling limits, we expect

$$f(\mathbf{B}) = \exp\left\{-\frac{1}{2}\lambda(1 - \cos\mathbf{B})\right\}; \quad \lambda \ll 1 \quad (4.10)$$

$$= \exp\left\{-\frac{1}{2}\sqrt{\lambda} \Gamma(\vec{0}) B^2\right\}; \quad \lambda \gg 1 \quad (4.11)$$

where it is understood that (4.11) represents a periodic gaussian.

Give the form of trial function (4.9) for the whole lattice, the optimal

single plaquette function $f(\mathbf{B})$ may be determined variationally by minimizing the energy functional

$$E[f] = \langle \phi | H | \phi \rangle / \langle \phi | \phi \rangle \quad (4.12)$$

By writing $f(\mathbf{B}) = \exp\left\{\frac{1}{2}u(\mathbf{B})\right\}$ and the expectation value of the kinetic energy in the corresponding Jackson-Feenberg form [27], the energy functional can be transformed to

$$\epsilon[f] = E[f]/N = \int_{-\pi}^{\pi} dA_1 dA_2 dA_3 dA_4 G(\mathbf{B}_q) \left[-\frac{1}{2}u''(\mathbf{B}_q) + \lambda(1 - \cos \mathbf{B}_q)\right], \quad (4.13)$$

where $\mathbf{B}_q$ is a particular plaquette formed by four links A_1 to A_4 and

$$G(\mathbf{B}_q) = f^2(\mathbf{B}_q) \frac{1}{\langle \phi | \phi \rangle} \int_{-\pi}^{\pi} dA_5 \dots dA_N \prod_{p \neq q} f^2(\mathbf{B}_p) \quad (4.14)$$

with normalization

$$\int_{-\pi}^{\pi} dA_1 dA_2 dA_3 dA_4 G(\mathbf{B}_q) = (2\pi)^3 \int_{-\pi}^{\pi} dB G(\mathbf{B}) = 1 \quad (4.15)$$

A technique in many-body theory which is useful in evaluating (4.14) is to develop a cluster expansion [27] in powers of the "bond" function

$$b(\mathbf{B}) = f^2(\mathbf{B}) - 1 \quad (4.16)$$

Such an expansion is clearly useful in the strong coupling limit where f is expected to be close to one and higher powers of b are correspondingly small. Substituting (4.16) in (4.14) yields

$$G(B)_q = f^2(B)_q \frac{1}{\langle \phi | \phi \rangle} \int_{-\pi}^{\pi} dA_1 \dots dA_N \left\{ 1 + \sum_p \prod_{p \neq q} b(B)_p + \sum_{p,k} \prod_{p \neq k \neq q} b(B)_p b(B)_k + \dots \right\} \quad (4.17)$$

We shall refer to an integrating link variable occurring in a product of bond functions as "exposed" if it occurs only once in the product. In addition, any bond function which contains at least one exposed link will also be defined to be "exposed." Since each bond function is periodic, any exposed bond can be immediately integrated via its exposed link to yield a constant. To prevent exposure, a link must be shared by at least two bonds. One can simply picture this requirement as the joining of each bond's plaquette at an edge. A product of bonds can therefore escape exposure only if its plaquettes form one or more closed surfaces. In the expansion (4.17), the only contributions to $G(B)_q$ which are not proportional to $f^2(B)_q$ correspond to connected, closed surfaces surrounding B_q . The first such term is a product of five bonds, whose plaquettes, together with B_q , form a closed cubic surface with unit volume. The next contribution occurs with nine bonds and is a closed rectangular surface with 2 units of volume, and so on.

Since the cluster expansion for $G(B)$ begins at such a high order, in the strong coupling limit, it is a reasonable approximation to retain only the zero order contribution. The normalization condition (4.15) can be satisfied by setting

$$G(B) = \frac{1}{(2\pi)^3} \frac{1}{\langle f | f \rangle} f^2(B) \quad (4.18)$$

Substituting this back into (4.13) gives

$$e[f] = \frac{\langle f | h | f \rangle}{\langle f | f \rangle} \quad (4.19)$$

with

$$h = -2 \frac{d^2}{db^2} + \lambda(1 - \cos b) \quad (4.20)$$

Thus in this approximation e is minimized by the ground state wavefunction of h . Since we have neglected the effect of other plaquettes on $G(B)$ by the approximation (4.18), it is natural that h is precisely the Hamiltonian for an isolated plaquette. Its ground state wavefunction (unnormalized) is the Mathieu function $ce_0\left(\frac{1}{2}B, -\lambda\right)$ with eigenvalue [29]

$$\epsilon_0 = \begin{cases} \lambda - \frac{1}{4}\lambda^2 + \frac{7}{256}\lambda^4 + \dots & \lambda \ll 1 \\ \sqrt{\lambda} - \frac{1}{8} + \dots & \lambda \gg 1 \end{cases} \quad (4.21)$$

and possesses the same strong and weak coupling behaviors of (4.10) and (4.11) with $1/2$ replacing $T(\vec{0})$. Thus a product of Mathieu functions may be expected to serve well as a trial function.

As a side remark, we note that in two dimensions, with free (i.e., no) boundary conditions, exposure is unavoidable. Thus every term in (4.17) is a constant and the entire sum simply collapses against the denominator $\langle \phi | \phi \rangle$ to give (4.18) exactly. Thus in two dimensions, for trial functions of the form (4.9), a product of Mathieu functions is the optimal variational wavefunction.

In the present work, rather than deal with numerical evaluation of Mathieu functions, it has been convenient to consider f of the form

$$f(B) = e^{-U(B)} \quad (4.22)$$

where $U(B)$ is even and periodic and is thus expandable in a cosine series. Keeping only the first two terms, we parametrize $U(B)$ by

$$U(B) = b[1 - \cos B] + \frac{1}{4}(a-b)[1 - \cos 2B], \quad (4.23)$$

where a and b are to be determined as functions of λ , by a Monte Carlo variational calculation. In the strong coupling limit, by comparison with (4.10), we expect

$$a = \frac{1}{2}\lambda, \quad (4.24)$$

$$b = \frac{1}{2}\lambda. \quad (4.25)$$

In the weak coupling limit, since the series expansion for U in power of B is

$$U(B) = \frac{1}{2!}aB^2 - \frac{1}{4!}(4a-3b)B^4 + \dots$$

Equation (4.11) is approximately reproduced by

$$a = \sqrt{\lambda} \quad r(\vec{0}) \quad (4.26)$$

and

$$b = \frac{4}{3}\sqrt{\lambda} \quad r(\vec{0}). \quad (4.27)$$

5. RESULTS

To demonstrate the utility of the present method, calculations were performed only on rather small lattices. With the exception of spatial Wilson loops, which were evaluated on a 6^3 lattice, all findings in this section are results of a 4^3 lattice computation.

At each value of λ , as a prelude to the main calculation, the trial function parameters a and b were determined variationally. The exact Monte Carlo evaluation of the variational energy (4.12) is accomplished by averaging $\phi^{-1} H \phi$ over a sequence of lattice configurations generated according to ϕ by the Metropolis [25] method. The procedure is precisely that of evaluating the trial energy at the initial step of the Guided Random Walk algorithm and, in practice, satisfactory statistics requires on the order of 10^3 configurations. The calculated λ -dependence of a and b are as shown on Fig. 1 and are in good agreement with the expectations (4.24) to (4.27). The resulting variational energies, plotted in Fig. 3 as triangles, reproduce the perturbative estimates in the strong coupling limit and match smoothly onto the harmonic results in the weak coupling regime. Thus from a variational perspective, ϕ is a reasonable approximation to ϕ_0 for all values of λ .

With ϕ defined by the variationally determined values of a and b , the Monte Carlo calculation with trial function guidance proceeds as outlined in Section 2.C. In the actual computation, we replicated according to the symmetrized weight $w(x_{n+1}, x_n) = \exp\left\{\Delta t \left[E - \frac{1}{2}\phi^{-1}(x_{n+1})H\phi(x_{n+1}) - \frac{1}{2}\phi^{-1}(x_n)H\phi(x_n)\right]\right\}$, as in Ref. 8. In retrospect, this is probably unnecessary. The inclusion of a variable step size $\Delta\tilde{t}(x)$, although in principle necessary, was found to have negligible effects on the final result (see Fig. 2 and accompanying discussion

below). Thus without significantly compromising the statistical accuracy of the calculation, we have performed computations in this work with a constant Δt . To determine the normalization energy, the constant E in the replication factor is changed after each time step to stabilize the average size of the ensemble at the initial value of N_p . In principle, frequent adjustment of E has the potential for introducing a bias in the final answer, which must be either corrected for explicitly or verified to be small. In conventional applications, with N_p on the order of hundreds, such a bias is found to be small. For example, in solving the one dimensional harmonic oscillator with unguided random walks, bias in E and $\langle x^2 \rangle$ are noticeable only for $N_p \lesssim 20$. As pointed out earlier in Section 2, to the extent that the adjustment in E is necessitated by the fluctuation in the ensemble population, the use of a guiding trial function greatly reduces this bias and enables the calculation to be performed with a smaller population. For lattice calculation, since the number of link variable increases rapidly with the lattice size, the use of a guiding trial wavefunction is essential to keep N_p as small as is possible. In the present calculation, although a smaller value may also be adequate, we use $N_p = 10$. (For $N_p = 10$, the unguided calculation yields noticeable bias in E_N and the plaquette energy $\langle 1 - \cos B \rangle$, as compared to the guided calculation. In a smaller 2^3 lattice, we have checked that this bias disappeared from the unguided calculation when $N_p \gtrsim 40$ and that results are then in agreement with the guided calculation with $N_p = 10$. Furthermore, in the guided calculation, no statistically significant differences were detected between computations with $N_p = 10$ and $N_p = 20$.)

In a typical run of 1000 to 2000 time steps (after discarding the first few hundreds which have not equilibrated to the ground state distribution), to assure statistical independence, observables were averaged over blocks of

20-50 steps. The resulting time step size dependence of E_T/N and E_N/N ($N = 3L^3$) at three selected values of λ are shown with solid lines of extrapolation on Fig. 2. The systematic trend from $\lambda = 1.8$ to $\lambda = 0.6$ clearly demonstrates the fact that in regions where ϕ is a good approximation to ϕ_0 , both statistical fluctuations and errors of Δt -extrapolation are simultaneously reduced. Also shown for comparison (dashed lines) is a $\lambda = 1.8$ calculation performed exactly as required by (2.19) with a variable step size $\tilde{\Delta t}(x)$. As the comparison makes clear, in the present calculation, neither the form of the replication factor nor the choice of the step size function has significant impact on the final, extrapolated result.

The energy per plaquette as a function of λ , obtained by Δt -extrapolation of E_N/N , is plotted on Fig. 3 as round dots. In the strong coupling regime, it differs little from the variational energy, denoted by triangles, or from the strong coupling expansion. In the weak coupling regime, as expected, it is significantly lower than the variational estimate, and is in excellent agreement with the expansion (4.6). This result provides convincing evidence that the stochastic evolution is capable of building up the off-diagonal order that is missing in the trial function, and that it converges to the proper ground state rather than becoming trapped in a metastable state corresponding to the wrong phase. It is not clear why the rather similar Greens Function Monte Carlo method does appear to become trapped in the wrong phase under similar circumstances. [9] Although it is not possible to study the nature of the phase transition without an appropriate scaling analysis [21], the smoothness of $e_0(\lambda)$ strongly suggests that a first order phase transition is unlikely.

The derivative $\partial \epsilon_0 / \partial \lambda$ may be calculated directly as an expectation value

$$\frac{\partial \epsilon_0}{\partial \lambda} = \frac{1}{N} \langle \phi_0 | \frac{\partial H}{\partial \lambda} | \phi_0 \rangle / \langle \phi_0 | \phi_0 \rangle \quad (4.1)$$

$$= \frac{1}{N} \langle \phi_0 | \sum_p (1 - \cos \pi p) | \phi_0 \rangle / \langle \phi_0 | \phi_0 \rangle \equiv \langle 1 - \cos \pi \rangle, \quad (4.2)$$

and will be referred to as the plaquette energy. Since $\langle 1 - \cos \pi \rangle$ is

essentially the potential energy, it can be accumulated easily during the same computer run to calculate ϵ_0 . The resulting Δt -dependence of $\langle 1 - \cos \pi \rangle$ is similar to ϵ_0 in that the slope decreases in regions where ϕ is close to ψ_0 .

In this case, since the individual variance no longer diminishes in proportion to $(\psi_0 - \phi)$, the extrapolation errors are substantial even for $\lambda \ll 1$. The plaquette energy, calculated as a function of λ according to the perturbative estimate (2.21), is shown on Fig. 4. We have checked that the plaquette energies calculated in this way for $\lambda < 1.1$ are consistent with results obtained by numerically differentiating $\epsilon_0(\lambda)$ of Fig. 3. However, the two data points for $\langle 1 - \cos \pi \rangle$ at $\lambda = 1.4$ ($\approx 0.28 \pm 0.01$) and at $\lambda = 1.8$ ($\approx 0.25 \pm 0.01$) are noticeably lower than the slope of $\epsilon_0(\lambda)$ obtained at $\lambda = 1.6$ by numerical differentiation, $\frac{d\epsilon_0}{d\lambda} = 0.315 \pm 0.015$. This latter result is in agreement with the weak coupling expansion, eq. (4.6) and is denoted by the cross in Fig. 9. Thus the plaquette energy has been slightly underestimated by the simple perturbative expression (2.21) in the weak coupling regime, indicating the need to use eq. (2.12) for precise evaluation of expectation values.

The plaquette energy's relatively rapid change of slope near $\lambda = 1$ is highly suggestive of a second order phase transition known from previous Lagrangian lattice calculation [21,22] and analytic study [20]. Since the present lattice is small, no sharp transition is really expected. Nevertheless, Fig. 4 clearly illustrates the utility of the present method in probing for nontrivial structures of a lattice gauge theory.

A gauge invariant order parameter which can distinguish the two phase of the lattice is the Wilson loop [28,16,18]. As a further application of the present method, we evaluated Wilson loops of various sizes on a 6^3 lattice at $\lambda = 0.9$ and at $\lambda = 2.0$. For the case of square loops with side S , the results are shown in Figs. 5 and 6, respectively. In this calculation, the variational parameters a and b were simply taken from Fig. 1 and were not determined anew for the 6^3 lattice. As expected, at $\lambda = 0.9$, Wilson loops show an area dependence which identifies the strong coupling regime as the confinement phase [28]. The normal QED phase without confinement is characterized by a perimeter dependence as shown for the case of $\lambda = 2.0$.

6. CONCLUDING REMARKS AND FUTURE PROSPECTS

In this work, we have presented a Monte Carlo method for solving Hamiltonian lattice gauge theories in which known physics can be incorporated by means of a trial function to guide the random walk process in sampling the ground state. To the extent that the trial function is a reasonable approximation to the ground state wavefunction, the method can be an efficient means of calculating the ground state energy and is therefore particularly suited for the study of first order phase transitions. For the latter case, the quantity of physical interest, the energy difference between the two phases (i.e., the Bag constant), is directly given by the algorithm.

We have demonstrated the utility of the method by solving the $U(1)$ lattice gauge theory in three spatial dimensions. Satisfactory results were obtained for the ground state energy and Wilson loops of various sizes. By essentially calculating the expectation value of the potential, we have further shown that even a second order phase transition can be unmasked by this technique.

By its use of a trial function, in contrast to existing techniques, the present method is potentially useful in clarifying the physics of the lattice ground state. Since the algorithm evaluates $\langle \phi | O | \phi \rangle$ initially and evolves it into $\langle \phi | O | \psi_0 \rangle$, the deviation of ϕ from ψ_0 relative to the operator O is directly measured. Thus one's physical understandings of ψ_0 , as embodied in ϕ , can be assessed quantitatively and modified accordingly. For this purpose, operators more sensitive to the ground state wavefunction than the Hamiltonian should be used.

The random walk algorithm is applicable to any Hamiltonian theories in which the kinetic part of the infinitesimal evolution matrix element is a

Gaussian distribution. In the case of $SU(N)$ Hamiltonian lattice gauge theory, the quadratic Casimir operator $E^a E^a$, which forms the kinetic part of the Hamiltonian, is simply proportional to the Laplace-Beltrami operator [30] $-\Delta = -g^{-1/2} \frac{\partial}{\partial x^i} \left\{ g^{1/2} g^{ij} \frac{\partial}{\partial x^j} \right\}$ on the $SU(N)$ group manifold parametrized by the coordinate system $\{x^i\}$ with metric $ds^2 = g_{ij} dx^i dx^j$. Up to a multiplicative constant, the metric tensor is given by [31] $g_{ij} = \text{tr}[(\partial U^i / \partial x^j)(\partial U^j / \partial x^i)]$, where U 's are elements of $SU(N)$. In this case, the lattice Hamiltonian can be regarded as describing a system of nonrelativistic interacting bosons on a (N^2-1) -dimensional curved space. The corresponding matrix element of the infinitesimal kinetic evolution operator is [30, 32]

$$\langle U' | \exp \left[-\frac{1}{2} \Delta t E^a E^a \right] | U \rangle \xrightarrow{\Delta t \rightarrow 0} (2\pi \Delta t)^{-(N^2-1)/2} \exp \left[-\frac{1}{2\Delta t} ds^2 \right], \quad (6.1)$$

where ds^2 is the metric element between U and U' . Since the group manifold is invariant under right (and left) translation, $\langle U' | \exp \left[-\frac{1}{2} \Delta t E^a E^a \right] | U \rangle = \langle U' U^{-1} | \exp \left[-\frac{1}{2} \Delta t E^a E^a \right] | U \rangle$, where I is the identity group element with coordinate $\vec{0}$. Hence, (6.1) can be sampled by setting $U' = \Delta U U$, where ΔU is a group element Gaussian distributed in distance from the origin with zero mean and variance $\langle ds^2 \rangle = \Delta t$. Since we can always choose a coordinate system locally Cartesian at the origin, the required ΔU is easily generated by a simple Cartesian Gaussian random walk from the origin. Thus with this slight generalization, the present Monte Carlo technique is also applicable to the study of $SU(N)$ lattice gauge theories.

Finally, the realization that lattice gauge theories are nonrelativistic many-boson systems defined on a group manifold suggests that they can also be studied by other established many-body theories such as the coupled-cluster expansion and variational techniques [27]. Such studies are interesting both

from the perspective of enriching the repertoire of many-body theories and as a potential source of insight into the lattice ground state.

ACKNOWLEDGMENTS

It is a pleasure to acknowledge fruitful discussions with D. Stump who is using the Green's function Monte Carlo method to solve the $U(1)$ Hamiltonian lattice gauge theory. We are grateful for the hospitality of the Institute for Theoretical Physics in Santa Barbara, where this project was conceived during the authors' participation in a program in Nuclear and Many-Body Physics. One of the authors (JWN) gratefully acknowledges fellowship support by the John Simon Guggenheim Memorial Foundation.

REFERENCES

1. M. H. Kalos, *Phys. Rev.* 128 (1962) 1791; *J. Comp. Phys.* 1 (1966) 257; *Phys. Rev.* A2 (1970) 250.
2. D. M. Ceperley and M. H. Kalos, *Monte Carlo Methods in Statistical Mechanics*, ed., K. Binder (Springer-Verlag, NY, 1979).
3. M. H. Kalos, D. Levesque, and L. Verlet, *Phys. Rev. A* 9, (1974) 2178.
4. P. A. Whitlock, M. H. Kalos, G. V. Chester, and D. M. Ceperley, *Phys. Rev. B* 19 (1979) 5598.
5. M. H. Kalos, M. A. Lee, P. A. Whitlock, and G. V. Chester, *Phys. Rev. B* 24 (1981) 115.
6. J. B. Anderson, *J. Chem. Phys.* 63 (1975) 1499; 65 (1976) 4121; 73 (1980) 3897; 74 (1981) 6307.
7. J. W. Moskowitz, K. E. Schmidt, M. E. Lee, and M. H. Kalos, *J. Chem. Phys.* 77 (1982) 349.
8. P. J. Reynolds, D. M. Ceperley, B. J. Alder, and W. A. Lester, *J. Chem. Phys.* 77 (1982) 5593.
9. D. Stump, talk presented at the 1982 Summer Workshop on Lattice Gauge Theory at the Institute for Theoretical Physics, Santa Barbara, CA; D. W. Heys and D. R. Stump, *Phys. Rev. D* 28 2067 (1983)
10. D. M. Ceperley and B. J. Alder, *Phys. Rev. Lett.* 45 (1980) 566.
11. D. M. Ceperley and B. J. Alder, *Physica B & C* 107 (1981) 875.
12. S. E. Koonin, *Nuclear Theory 1981*, ed., G. F. Bertsch (World Scientific, 1981).
13. J. G. Zabolitzky and M. H. Kalos, *Nucl. Phys.* A356 (1981) 114.
14. J. W. Negele, *Proc. Int. Symp. on Time-Dependent Hartree-Fock and Beyond*, Lecture Notes in Physics, Vol. 171, eds., K. Goeke and P. G. Reinhardt (Springer-Verlag, NY, 1979).
15. B. D. Serot, K. E. Koonin, and J. W. Negele, *Phys. Rev. C* (to be published).
16. K. Kogut and L. Susskind, *Phys. Rev. D* 11 (1975) 395.
17. A. M. Polyakov, *Phys. Lett.* 72B (1978) 477.
18. E. Fradkin and L. Susskind, *Phys. Rev. D* 17 (1978) 2637.
19. S. D. Drell, H. R. Quinn, B. Svetitsky, and M. Weinstein, *Phys. Rev. D* 19 (1979) 619.
20. A. H. Guth, *Phys. Rev. D* 21 (1980) 2291.
21. B. Lautrup and M. Nauenberg, *Phys. Lett.* 95B (1980) 63.
22. K. J. M. Moriarty, *Phys. Rev. D* 25 (1982) 2185.
23. A. Einstein, *Investigations on the Theory of Brownian Movement* (Dover, 1956).
24. J. W. Negele, *Proceedings, III International Conf. on Recent Progress in Many Body Theories*, Altenberg, Germany (Springer-Verlag, 1983 [in press]).
25. N. Metropolis, A. Rosenbluth, M. Rosenbluth, A. Teller, and E. Teller, *J. Chem. Phys.* 21 (1953) 1087.
26. J. M. Hamersley and D. C. Handscomb, *Monte Carlo Method* (Methuen, London, 1964).
27. G. Ripka, *Phys. Rept.* 56 (1979) 1.
28. K. Wilson, *Phys. Rev. D* 10 (1974) 2445.
29. M. Abramowitz and I. A. Stegun, eds., *Handbook of Mathematical Functions*, National Bureau of Standards (1964).
30. N. E. Hurt, *Geometric Quantization in Action* (D. Reidel, 1983).
31. J. M. Normand, *A Lie Group: Rotations in Quantum Physics* (North Holland, Amsterdam, 1980).
32. J. S. Dowker, *Ann. Physics* 62 (1971) 361.

FIGURE CAPTIONS

Fig. 1. The trial function parameters a and b as determined by a Monte Carlo variational calculation for a 4^3 lattice. The straight and curved line indicate expected behaviors in the strong and weak coupling regime. The error bars correspond to grid sizes used to locate the energy minimum.

Fig. 2. The calculated ground state energy per plaquette for a 4^3 lattice at selected values of λ as a function of time step size Δt . The normalization and trial energies are indicated by the square and round symbols respectively. The error bars are calculated from block average of 20, 20, and 50 time steps for the case of $\lambda = 0.6$, 1.0, and 1.8, respectively. The solid straight lines are possible extrapolations drawn to guide the eye. Results connected by dashed lines are calculated according to (2.19) using a variable time step and an asymmetric replication factor.

Fig. 3. The ground state energy per plaquette as a function of λ . The solid lines denote the strong and weak coupling expansions, eqs. 4.2 and 4.6. The triangles and dots denote variational and exact Monte Carlo results for a 4^3 lattice. The statistical errors in both cases are too small to be shown.

Fig. 4. The plaquette energy, corresponding to $\partial \epsilon_0 / \partial \lambda$, as a function of λ for a 4^3 lattice. The solid lines denote the strong and weak coupling expansions obtained from eqs. 4.2 and 4.6. The Monte Carlo results using the perturbative expression eq. (2.21) are denoted by the solid points, connected by dashed

lines with error bars calculated by combining statistical errors entering into the Δt -extrapolation. The cross at $\lambda = 1.6$ indicates the slope obtained by differentiating the ground state energy as explained in the text.

Fig. 5. Calculated expectation values of square spatial Wilson loops as a function of area S^2 for a 6^3 lattice at $\lambda = 0.9$.

Fig. 6. Calculated expectation values of square spatial Wilson loops as a function of side length S for a 6^3 lattice at $\lambda = 2.0$.

Fig. 1

Fig. 2

Fig. 3

Fig. 4

Fig. 5

Fig. 6