PHYSICAL REVIEW B **70**, 193101 (2004)

Truncated-determinant diagrammatic Monte Carlo for fermions with contact interaction

Evgueni Bourovski, Nikolay Prokof'ev, and Boris Svistunov
Physics Department, University of Massachusetts, Amherst, Massachusetts 01003, USA
and Russian Research Center "Kurchatov Institute," 123182 Moscow, Russia
(Received 1 July 2004; published 4 November 2004)

For some models of interacting fermions the known solution to the notorious sign problem in Monte Carlo (MC) simulations is to work with macroscopic fermionic determinants; the price, however, is a macroscopic scaling of the numerical effort spent on elementary local updates. We find that the *ratio* of two macroscopic determinants can be found with any desired accuracy by considering truncated (local in space and time) matices. In this respect, MC for interacting fermionic systems becomes similar to that for the sign-problem-free bosonic systems with system-size independent update cost. We demonstrate the utility of the truncated-determinant method by simulating the attractive Hubbard model within the MC scheme based on partially summed Feynman diagrams. We conjecture that similar approach may be useful in other implementations of the sign-free determinant schemes.

DOI: 10.1103/PhysRevB.70.193101 PACS number(s): 02.70.Ss, 03.75.Ss, 71.10.Fd

Monte Carlo (MC) methods are a unique tool for studying large interacting systems. The most severe limitation on their applicability is imposed by the so-called sign problem (SP) when relevant contributions to statistics alternate in sign and almost exactly compensate each other in the final answer.[1] Frustrating interactions and anticommutation relations for fermion operators are typically at the origin of the sign problem. In this paper, we address the case when quantum statistics is nonpositive only because of the fermion exchange cycles.

One solution to the fermion SP is offered by the determinant Monte Carlo (DetMC) (see, e.g., Ref. 2). The idea is that all contributions to the many-body statistics obtained by exchanging fermion places (in a certain representation) can be written as a product of two determinants—for spin-up and spin-down species—and in cases when the two real determinants coincide the result is positive definite. In Metropolis-type algorithms,[3] MC updates are accepted with probabilities proportional to the ratio of final and initial configuration weights; in DetMC the corresponding acceptance ratio R is based on the ratio of large determinants. Unfortunately, calculating determinants ratio for macroscopically large matrices is very expensive numerically: even with tricks involving the Hubbard-Stratonovich tranformation the algorithm proposed by Blankenbecler, Scalapino, and Sugar[4,5] still requires L^{2D} operations per update for a D-dimensional system with L lattice points per dimension. The same scaling is true for the continuous-time scheme.[6] In contrast, for bosonic systems with local interactions the number of operations per update is small and system-size independent, i.e., they can be simulated L^{2D} times faster.

Since the bottleneck of DetMC is the calculation of R, one may question the paradigm of calculating it "exactly:" In any case, computer operations always involve systematic round-off errors. The other example is provided by (pseudo)random number generators—they are *always* imperfect and result in systematic errors equivalent to small errors in R which, however, remain practically undetectable (for good generators) in final results. The heuristic explanation of why small errors in R do not ruin the simulation is as follows. The Metropolis algorithm is a scheme with strong relaxation towards equilibrium distribution, and local configuration updates may be viewed as the result of dissipative coupling to the thermal bath (this picture is often used to model dissipative kinetics[1]). Uncontrolled errors in R may then be regarded as a small stochastic noise in the relaxational dynamics. As such, it only slightly modifies the equilibrium state and its properties. This is a standard argument in the linear response theory.

It seems natural then to suggest that if the goal is to simulate the result with n-digit accuracy, there is no need to calculate the acceptance ratio with accuracy much higher than n digits. Often, we simply ignore this issue because getting R with machine precision does not cost any extra CPU time. In determinant methods, however, there is a potential of huge efficiency gains if approximate values of R can be calculated much faster. We demonstrate the feasibility of this approach by showing that the ratio of two macroscopic determinants can be found with high accuracy by considering truncated matrices dealing only with the local (in space-imaginary time) structure of the configuration space. The computational cost of updates in the corresponding "truncated-determinant" scheme is system-size independent—an efficiency increase $\propto L^{2D}$ for large L.

In what follows, we discuss the solution of the Hubbard model for fermions within a simple diagrammatic MC scheme based on Feynman diagrams partially summed over fermion propagator permutations.[6,7] The resulting diagram weight is the square of the determinant composed of finite-temperature fermion propagators. We explain how the determinant ratio for local updates may be calculated using truncated matrices, and demonstrate the feasibility of the proposed approach. We also show that going to larger system sizes has little effect on the scheme performance. Finally, we conjecture that large efficiency gains are expected in other sign-problem free DetMC schemes, e.g., in lattice QCD simulations with quark fields.[8] It is also worth noting that in the diagrammatic DetMC scheme the update cost does not depend directly on the lattice period, which is a big advantage for simulations of dilute systems. By "dilute" we mean

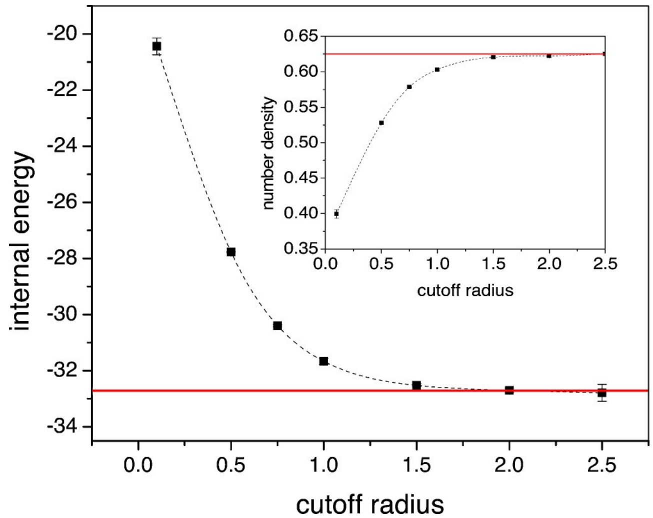

FIG. 1. Energy and density dependence on the imaginary time cutoff for the Hubbard model for $L=4$. Dots represent the MC data, and solid lines are the exact diagonalization results for 10 particles (Ref. 11).

dilute with respect to the lattice, but not necessarily with respect to the interaction; the latter can be effectively large due to a resonance on a (quasi) bound state. Correspondingly, the new scheme is very promising for the study of ultra cold fermionic systems in the regime of strong Feshbach resonant interaction, including the crossover from the Bardeen-Cooper-Schrieffer pairing to the Bose-Einstein condensation of molecules.[16]

Model and method. We consider interacting lattice fermions with the Hubbard Hamiltonian $H=H_0+H_1$:

$$H_0 = \sum_{\mathbf{k}\sigma} (\epsilon_k - \mu) c^{\dagger}_{\mathbf{k}\sigma} c_{\mathbf{k}\sigma}, \; H_1 = U \sum_x n_{x\uparrow} n_{x\downarrow}, \tag{1}$$

where $c^{\dagger}_{\mathbf{x}\sigma}$ is the fermion creation operator, $n_{\mathbf{x}\sigma}=c^{\dagger}_{\mathbf{x}\sigma}c_{\mathbf{x}\sigma}$, $\sigma=\uparrow,\downarrow$ is the spin index, $\mathbf{x}$ runs over the L^D points of the simple cubic lattice, $\mathbf{k}$ runs over the corresponding Brillouin zone, $\epsilon_{\mathbf{k}}=-2t\sum_{\alpha=1}^{D}\cos k_\alpha a$ is the tight-binding dispersion law, and μ is the chemical potential. For definiteness and numerical tests, we confine ourselves to the $D=2$ spacial lattice with periodic boundary conditions. We use the hopping amplitude, t, and lattice constant a as units of energy and distance, respectively.

Following Refs. 6 and 7 we start with writing the statistical operator in the interaction representation

$$e^{-\beta H} = e^{-\beta H_0} \mathcal{T}_\tau \exp\left\{ - \int_0^\beta H_1(\tau) d\tau \right\}, \tag{2}$$

where $H_1(\tau)=e^{H_0\tau}H_1 e^{-H_0\tau}$ and $\mathcal{T}_\tau$ is the time ordering operator, and expanding it in powers of H_1:

$$Z = \sum_{p=0}^{\infty} (-U)^p \sum_{\mathbf{x}_1\cdots\mathbf{x}_p} \int_{\tau_1<\tau_2<\cdots<\tau_p<\beta} \left(\prod_{i=1}^{p} d\tau_i \right) \times \mathrm{Tr}\left[e^{-\beta H_0} \prod_{i=1}^{p} c^{\dagger}_{\uparrow}(\mathbf{x}_i\tau_i) c_{\uparrow}(\mathbf{x}_i\tau_i) c^{\dagger}_{\downarrow}(\mathbf{x}_i\tau_i) c_{\downarrow}(\mathbf{x}_i\tau_i) \right]. \tag{3}$$

This expansion for the partition function generates standard Feynman diagrams.[9] Graphically, each term is a set of four-point vertices with two incoming (spin up and spin down), and two outgoing (spin up and spin down) lines which connect vertices. Each line is associated with the imaginary time fermion propagator $G_\sigma(\mathbf{x}_i-\mathbf{x}_j,\tau_i-\tau_j;\mu,\beta) = -\mathrm{Tr}[\mathcal{T}_\tau e^{-\beta H_0} c_\sigma(\mathbf{x}_i\tau_i) c^{\dagger}_\sigma(\mathbf{x}_j\tau_j)]$.

A straightforward MC sampling of diagrammatic series would be impossible because of the sign problem. However, if for a given configuration of p vertices, $\mathcal{S}_p=\{(\mathbf{x}_j,\tau_j),\ j=1,\ldots,p\}$, one sums over all $(p!)^2$ ways of connecting them by propagators, then the result can be written as a *product of two determinants*, one for spin up, and another for spin down (see, e.g., Ref. 2). The differential weight of the vertex configuration (or vertex diagram) is then

$$d\mathcal{P}(\mathcal{S}_p) = (-U)^p \det \mathbf{A}^{\uparrow}(\mathcal{S}_p) \det \mathbf{A}^{\downarrow}(\mathcal{S}_p) \prod_{i=1}^{p} d\tau_i, \tag{4}$$

where $\mathbf{A}^\sigma(\mathcal{S}_p)$ are $p\times p$ matrices $A^\sigma_{ij}=G_\sigma(\mathbf{x}_i-\mathbf{x}_j,\tau_i-\tau_j)$. For equal number of up and down particles, $\det \mathbf{A}^{\uparrow} \det \mathbf{A}^{\downarrow}=[\det \mathbf{A}]^2$, and negative U the vertex diagram weight is always positive. (At half filling, $n_\uparrow+n_\downarrow=1$, the sign of U changes when hole representation is used for one of the spin components, so this scheme may be also used for the repulsive Hubbard model.)

MC sampling in the vertex configuration space $(p,\mathcal{S}_p)$ can be performed by standard MC rules (see, e.g., Refs. 7 and 10) using just one pair of complementary updates $\mathcal{D}$ and $\mathcal{C}$: in $\mathcal{D}$ one selects at random one of the vertices and suggests to delete it from the configuration; in $\mathcal{C}$ an additional vertex is suggested to be inserted at some point randomly selected in the space-time box $\beta\times L^D$. These updates decrease/increase the rank of $\mathbf{A}$ by one. The acceptance ratio for the $\mathcal{D}/\mathcal{C}$ pair of updates is then based on the ratio of two determinants

$$R_p = \det \mathbf{A}(\mathcal{S}_{p+1}) / \det \mathbf{A}(\mathcal{S}_p), \tag{5}$$

where $\mathcal{S}_{p+1}=\{\mathcal{S}_p,(\mathbf{x}_{p+1},\tau_{p+1})\}$ (we omit the spin index for brevity).

The bottleneck of this simple scheme is in evaluating R_p when p is macroscopically large. The typical number of vertices is determined by the number of particles, interaction strength, and inverse temperature as $p\propto N\beta U$. The truncated-determinant idea is to calculate R_p much faster at the expense of accuracy using the following conjecture originating from physical, rather than mathematical, arguments. The vertex configuration represents a sequence of virtual particle collisions in the many-body system, and it is likely that *local* changes in its structure depend only on the immediate neighborhood of the updated region. (We note that the idea of employing the local nature of the fermion-boson coupling has been used in Ref. 4, but it has not been extended to the fermionic determinant.) Quantitatively, we define a norm $\|\ldots\|$ or a distance, between vertices in space-time (several choices are discussed below), and construct a truncated vertex configuration $\mathcal{S}_p^{(l)}$ such that all points in $\mathcal{S}_p^{(l)}$ satisfy

$$\|(\mathbf{x}_j, \tau_j) - (\mathbf{x}_{p+1}, \tau_{p+1})\| \leq l. \quad (6)$$

Correspondingly, $\mathcal{S}_{p+1}^{(l)} = \{\mathcal{S}_p^{(l)}, (\mathbf{x}_{p+1}\tau_{p+1})\}$. We may now use truncated configurations to calculate the ratio (5) approximately as

$$R_p^{(l)} = \det \mathbf{A}(\mathcal{S}_{p+1}^{(l)}) / \det \mathbf{A}(\mathcal{S}_p^{(l)}). \quad (7)$$

Clearly, when $l \to L$ we recover the exact ratio. Our conjecture is then that $R_p^{(l)}$ quickly converges to R_p and there exists a healing length in the $(\mathbf{x}, \tau)$ space characterizing this convergence. If this is the case, then l may be considered as a microscopic (system-size independent) parameter controlling the accuracy and efficiency of simulation.

The proper choice of the norm $\|\cdots\|$ depends on system parameters. In the strongly correlated case the natural units of distance and time are provided by the Fermi momentum k_F and Fermi energy ϵ_F. One possibility is then

$$\|(\mathbf{x}, \tau) - (\mathbf{x}', \tau')\| = \sqrt{k_F^2(\mathbf{x} - \mathbf{x}')^2 + \epsilon_F^2(\tau - \tau')^2}. \quad (8)$$

Geometrically, this measure results in a set of vertices $\mathcal{S}_p^{(l)}$ inside the space-time ellipsoid centered at $(\mathbf{x}_{p+1}, \tau_{p+1})$. For dense systems $\|(\mathbf{x}, \tau) - (\mathbf{x}', \tau')\| = \max\{|\mathbf{x} - \mathbf{x}'|, |\tau - \tau'|\}$ is equally appropriate and our data for the largest system were obtained using this measure. At temperatures comparable to ϵ_F one may account for all vertices in the $\hat{\tau}$ direction and simply write $\|(\mathbf{x}, \tau) - (\mathbf{x}', \tau')\|_{\mathrm{cyl}} = k_F|\mathbf{x} - \mathbf{x}'|$. The corresponding geometrical figure is a β cylinder. Similarly, in small systems one may consider truncating configurations only in time direction.

Numerical results and discussion. Our tests of the truncated-determinant scheme were done for the attractive Hubbard model with $U = -4$, $\mu = -2$, $\beta = 10$, and periodic boundary conditions. First, we simulated a small $L^2 = 4^2$ cluster for which the ground-state energy (of 10 particles) is known from exact diagonalization studies.[11] Since the spatial dimension is so small we truncate vertex configurations only in the imaginary time direction. In Fig. 1 we show how the result for energy converges to the exact value. We stress, that at all stages of the MC simulation we never even write the full configuration determinant, which rank is about 2.5 times larger than typical values of p for the cutoff radius 2.

In Fig. 2 we present our data for the $L^2 = 100^2$ system—now, using determinant truncation both in space and in time directions, Eq. (8). Remarkably, the convergence is achieved around the same value of the truncation radius, which proves that the computational cost per update is not subject to macroscopic scaling. It is instructive to see how data convergence for energy correlates with the typical errors introduced by the approximate calculation of R_p. In Fig. 3 we show examples of R_p dependence on the truncation radius for a number of randomly selected MC configurations. Clearly, quite large fluctuations in R_p are statistically "averaged out" in the final result for energy. We are not aware of any other method capable of simulating fermionic systems of comparable size.

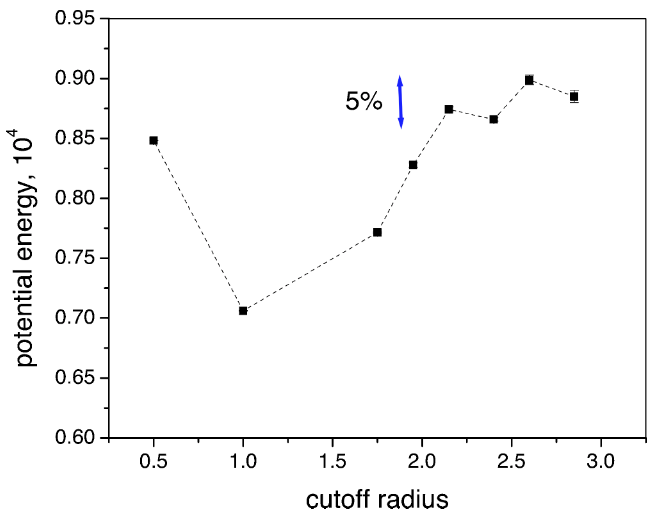

FIG. 2. Potential energy dependence on the truncation radius for the Hubbard model for $L = 100$.

Apart from the Hubbard model tests, we have also verified that the use of truncated-determinants for randomly seeded vertex configurations works as nicely to speed up the calculation of R_p.

The benchmark DetMC method by Blankenbecler, Scalapino, and Sugar[4,5] (BSS) is based on the Trotter-Suzuki imaginary time slicing, and the Hubbard-Stratonovich transformation.[12] The rank of the matrix used in the calculation of the acceptance ratio equals the number of lattice sites L^D and $\sim L^{2D}$ operations are required to find its determinant. The necessity of handling large matrices, although made possible by this method, requires both elaborate finite-size scaling analysis of MC data,[13] and special efforts for the calculation stabilization at low temperatures.[2] The contour-distortion stabilization techniques (see, e.g., Refs. 14 and 15) help to alleviate the sign problem, but still suffer from the severe scaling of the computational cost per update. More recently, Rombouts, Heide, and Jachowicz[6] improved the

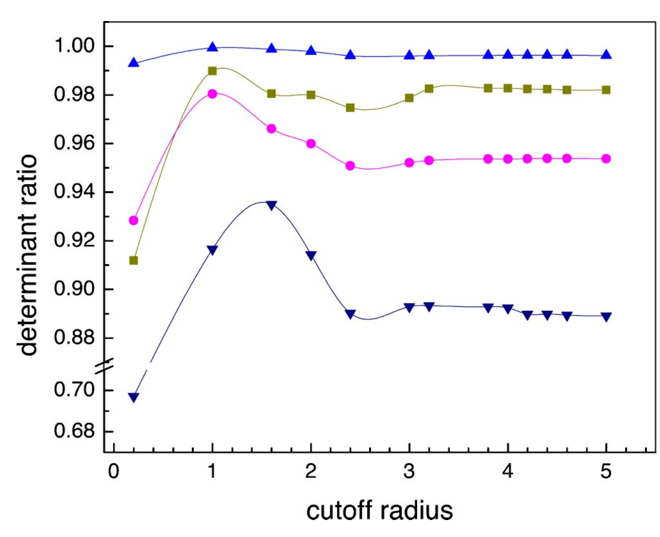

FIG. 3. (Color online) Determinant ratios $R_p^{(l)}$ as functions of l for randomly selected MC configurations for the $L = 100$ system.

BSS scheme by formulating it in the τ continuum. It is easy to directly compare this scheme with ours because the starting point is exactly the same—the expansion of the statistical operator in powers of U. Rombouts *et al.* used the auxiliary Ising variables to decompose four-point vertices into the sums of arrived at the number of operations for performing one update scaling as L^{2D}. At this point we notice that while we work with the same vertex configuration structure, in our scheme there is no extra summation over the auxiliary variables, and the calculation of the acceptance ratio is system-size independent.

Recently, a substantial improvement in lattice QCD simulations has been achieved by including quark-loop effects (see, e.g., Ref. 8). After the quarks are integrated out, their effects are described by the macroscopic positive definite determinant $\det \mathbf{A}(U)$, where U is the configuration of gluon matrices. The conjecture is that the ratio $\det \mathbf{A}(U)/\det \mathbf{A}(U')$ for local updates of gluon matrices $U \to U'$ can be calculated by accounting only for the immediate neighborhood of the updated lattice bond.

We doubt that the truncated-determinant schemes will help to speed up simulations with the sign problem. If the average configuration sign is small, the answer is determined by small differences between the sign-positive and sign-negative contributions, i.e., each contribution has to be calculated to much higher accuracy than would be sufficient in the positive definite case. As a result, higher and higher precision is required for R_p and the advantages of our approach quickly vanish.

We acknowledge support by the NSF Grant No. ITR-0426881.

[1] K. Binder and D. P. Landau, *A Guide to Monte Carlo Simulations in Statistical Physics* (Cambridge University Press, Cambridge, 2000).

[2] E. Y. Loh and J. E. Gubernatis, in *Electronic Phase Transitions*, edited by W. Hanke and Y. V. Kopaev (Elsevier Science Publishers, New York, 1992).

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

[4] D. J. Scalapino and R. L. Sugar, Phys. Rev. Lett. **46**, 519 (1981).

[5] R. Blankenbecler, D. J. Scalapino, and R. L. Sugar, Phys. Rev. D **24**, 2278 (1981).

[6] S. M. A. Rombouts, K. Heyde, and N. Jachowicz, Phys. Rev. Lett. **82**, 4155 (1999).

[7] A. N. Rubtsov, cond-mat/0302228 (unpublished).

[8] C. DeTar and S. Gottlieb, Phys. Today **57**, 45 (2004).

[9] A. L. Fetter and J. D. Walecka, *Quantum Theory of Many-particle Systems* (McGraw-Hill, New York, 1971).

[10] N. V. Prokof'ev and B. V. Svistunov, Phys. Rev. Lett. **81**, 2514 (1998).

[11] T. Husslein, W. Fettes, and I. Morgenstern, Int. J. Mod. Phys. C **8**, 397 (1997); cond-mat/9705026 (unpublished).

[12] J. E. Hirsch, Phys. Rev. B **28**, 4059 (1983).

[13] A. Sewer, X. Zotos, and H. Beck, Phys. Rev. B **66**, 140504 (2002).

[14] D. Charutz, D. Neuhauser, and N. Rom, Chem. Phys. Lett. **270**, 382 (1997).

[15] S. A. Baeurle, Phys. Rev. Lett. **89**, 080602 (2002).

[16] Barbara Goss Levi, Phys. Today **56**, 18 (2003), and references therein; for recent results see D. S. Petrov, C. Salomon, and G. V. Shlyapnikov, cond-mat/0309010 (unpublished); M. Greiner, C. A. Regal, and Deborah S. Jin, Phys. Rev. Lett. **92**, 040403 (2004); S. Jochim *et al.*, Science **32**, 2101 (2003); M. W. Zwierlein *et al.*, Phys. Rev. Lett. **91**, 250401 (2003).