LNF-94/002

Microcanonical Fermionic Average Method for Monte Carlo Simulations of Lattice Gauge Theories with Dynamical Fermions

V. Azcoiti, V. Laliena, X.Q. Luo, C.E. Piedrafita, G. Di Carlo, A. Galante, A.F. Grillo, L.A. Fernàndez, A. Vladikas

Physical Review D 48, 1, 402-416, (1993)

Microcanonical fermionic average method for Monte Carlo simulations of lattice gauge theories with dynamical fermions

V. Azcoiti, V. Laliena, X. Q. Luo, and C. E. Piedrafita
Departamento de Física Teórica, Facultad de Ciencias, Universidad de Zaragoza, 50009 Zaragoza, Spain

G. Di Carlo, A. Galante,* and A. F. Grillo
Istituto Nazionale di Fisica Nucleare, Laboratori Nazionali di Frascati, P.O. Box 13, Frascati, Italy

L. A. Fernández
Departamento de Física Teórica I, Universidad Complutense de Madrid, 28040 Madrid, Spain

A. Vladikas
Istituto Nazionale di Fisica Nucleare (INFN), Sezione di Roma I, Universitá di Roma I, La Sapienza, Piazzale Aldo Moro 2, 00185 Roma, Italy

(Received 25 January 1993)

We present a comprehensive exposition of a method for performing numerical simulations of lattice gauge theories with dynamical fermions. Its main aspects have been presented elsewhere. This work is a systematic study of the feasibility of the method, which amounts to separating the evaluation of the fermionic determinant from the generation of gauge configurations through a microcanonical process. The main advantage consists in the fact that the parts of the simulation which are most computer intensive must not be repeated when varying the parameters of the theory. Moreover, we achieve good control over critical slowing down, since the configurations over which the determinant is measured are always very well decorrelated; in addition, the actual implementation of the method allows us to perform simulations at exactly zero fermion mass. We relate the numerical feasibility of this approach to an expansion in the number of flavors; the criteria for its convergence are analyzed both theoretically and in connection with physical problems. On more speculative grounds, we argue that the origin of the applicability of the method stems from the nonlocality of the theory under consideration.

PACS number(s): 11.15.Ha, 02.70.Lq, 12.20.Ds

I. INTRODUCTION

There has been considerable progress in the last few years in the field of numerical simulations of field theories on the lattice [1]. The main constraint, which characterizes the field from its birth, has been the enormous amount of computer resources which are required in order to have a reliable simulation with small statistical and systematic errors. With the development of more powerful computers [2, 3] and algorithms [4], this aim is gradually being achieved and the quality of the measurements of various physical quantities is being improved. Data improvement involves a twofold procedure. On the one hand, statistical errors must be reduced by increasing the number of field configurations which are generated by the simulation method. On the other hand, the scaling region has to be approached by increasing the lattice size and tuning the field theory's parameters near their critical values. Only then is the discretized (lattice) field theory approximating its continuum counterpart. We are chiefly interested in simulating gauge theories, such as QED and QCD in four space-time dimensions. Thus, the simultaneous satisfaction of the above requirements (i.e., production of a large number of configurations on large four-dimensional lattices) has proved a highly nontrivial task for present day computers.

Historically speaking, the strategy adopted was to start with the simulation of reduced versions of a given lattice field theory and gradually work towards the complete theory, with the aid of increasingly more powerful computers and numerical methods. Thus, the first generation of computer simulations dealt with pure gauge theories. As these models involve only the computation of a local action with complex degrees of freedom (the gauge bosons), this proved a relatively easy task. The next step was to include the effect of fermion fields. The fermionic action consists of a highly nonlocal determinant term, which is extremely costly to calculate numerically. Consequently, the determinant term is habitually set to unity, in what is known as the quenched approximation [5]. In the quenched approximation the theory describes the interactions of gauge bosons with valence fermions; internal fermion loops are suppressed by the quenching.

*Also at Dipartimento di Fisica dell'Universitá dell'Aquila, L'Aquila 67100, Italy.

Quenched fermionic simulations may be regarded as the second generation of lattice field theory computations. The third generation is the simulation of the complete theory, including dynamical fermions; i.e., the effect of the fermionic determinant is fully taken into account. At present, quenched simulations are in a satisfactory state, although a lot remains to be done [6]. On the contrary, unquenched simulations are still at an embryonic level, although a lot of progress has been achieved [6].

The aim of this paper is to present a full exposition of a new method for the simulation of dynamical fermions, describe its implementation and the technicalities involved, discuss its shortcomings and advantages in detail, and present some physics results obtained with it. Several important aspects of these ideas have already been given in Refs. [7–11]. In these works, the emphasis was mainly on the properties concerning the critical behavior of compact and noncompact lattice QED. In the present paper, we are concentrating on the novelties of the method we introduced for the simulation of dynamical fermions. In particular, we address the question of how the feasibility of the method may be related to the nonlocality of the action. Moreover, a systematic discussion of the approximations and their influence on the physical results is presented. In doing so, different physical models have been used as examples (QED, Ising, etc.).

The main efforts for the simulation of dynamical fermions have been in the direction of algorithm development. There has been a lot of progress; the earlier proposals involved computational costs of $O(V^2)$ [12] [the pseudofermion algorithm in practical application being $O(V)$ but at the price of introducing systematic approximations], whereas the more recent algorithms have a theoretical cost of almost $O(V)$ (for example, the hybrid algorithm [13] or hybrid Monte Carlo algorithm [14]). However, these costs are theoretical order-of-magnitude estimates of the dependence of the computation of the fermionic determinant on the lattice volume V. In realistic calculations near criticality, the cost is mainly determined by critical slowing down, for which some remedies have been proposed such as Fourier acceleration [15] and lower-upper decomposition of the fermion matrix [16].

Another factor affecting the computational cost of the simulations is the relatively large number of parameters that characterize a field theory. These are normally the fermion masses m, the inverse gauge coupling β, and the number of flavors n_f. In order to tackle a typical lattice problem, the calculation must be repeated for, say, M mass values, B gauge couplings, and sometimes for F different numbers of flavors. Thus, the real cost of the computation is $C_1 C_2 V^p$, where C_1 is a factor that depends on the algorithm and the value of theory's parameters, p is the volume dependence already discussed above (both C_1 and p include critical slowing down effects), and $C_2 = M \times B \times F$ is the repetition factor. This factor is determined by the number of times for which we have to repeat the calculation in order to achieve a satisfactory understanding of the physical properties under study. In particular, the fermion mass enters crucially, in the sense that the continuum limit (at least for the study of the chiral limit) has to be obtained approaching zero bare fermion mass, which, with traditional methods, becomes very costly.

The considerable progress we have sketched above involves a reduction of C_1 and p by the development of fast algorithms and acceleration techniques, but does not improve C_2. Moreover, all the above methods imply measurements of the fermionic determinant at each upgrade of the gauge configuration. We propose, instead, a method which should avoid both these problems. First, it reduces C_2 by separating the simulation into two parts, as will be explained below. Second, we improve upon C_1 in the sense that a good control of critical slowing down effects of the fermion dynamics is easily implementable.

The method we will describe is based on reexpressing the theory in terms of an "averaged" determinant, which is a function of the system's pure gauge energy and of the fermion mass, and independent on β. All other dependencies are averaged out [7, 8]. In Sec. II we define this "averaged" determinant, which we call the effective fermionic action. The idea is that it may then be calculated numerically for a wide energy range, either at a fixed mass or by means of the determination of all the zero mass eigenvalues of the fermionic matrix. In the latter case the determinant is known for all mass values. As the average determinant does not depend on β, this calculation, which is the costliest, is performed only once in the second case or once for each mass in the first case. We may then use this result in a standard simulation of the reexpressed theory; the determinant is now a known function of the energy and the pure gauge part is not costly to simulate. The repetition factor is essentially reduced to $C_2 = M$ or $C_2 = 1$. In the latter case ($C_2 = 1$) we stress that the computation is feasible for every value of the fermion mass, zero included. The number of flavors enters trivially in this formulation and does not increase the computational cost significantly.

This method has been applied so far to the simulation of Abelian models, both compact and noncompact. The results of [7–10] refer to four-dimensional theories, and work is in progress in simulations in two and three dimensions [11]. In all the cases analyzed, our method has given results in agreement with more traditional approaches, at a small fraction of their computational cost [17].

While in principle it is possible to measure the determinant without approximations, for small fermion masses we might need a huge amount of statistics in order to measure accurately the effective fermionic action, as we will argue in Sec. III. This depends on the form of the probability distribution of the fermionic determinant. In order to overcome this problem, we introduce in Sec. IV an approximation [9]: The effective fermionic action, obtained from our "averaged" determinant, is expanded in powers of the number of flavors n_f; the result is essentially a cumulant expansion which we truncate after the first two or three terms.

This will be shown to be a satisfactory approximation in that the truncated terms matter little. In practice, we have seen that for $n_f \leq 4$ our cumulant expansion manages to simulate the theory without any significant loss of the dynamical fermion effects [7–9]. In theory, the

method may be viewed as interpolating between quenching and unquenching, having much of the advantages of quenched simulations in terms of CPU utilization, but at the same time giving results indistinguishable from traditional unquenched methods. In Sec. V we show how our method may be implemented in the calculation of physical observables, such as the chiral condensate.

Section VI summarizes some results of physical interest obtained with this method.

In Sec. VII we discuss the reliability of the truncated cumulant expansion in connection with the theory's structure (in particular the range of the interaction) and the dimensionality of space-time. The Ising model and the mixed compact-noncompact QED are used as case studies. We also discuss the strong resemblance of the lowest-order cumulant expansion to the mean field approximation.

Section VIII deals with some technical points. We describe the implementation both of the microcanonical and canonical simulations, and the Lanczos algorithm which we use in the calculation of the fermionic determinant of a given gauge field configuration. We discuss the practical implementation of this method in transputer networks in Sec. IX

II. EFFECTIVE FERMIONIC ACTION

In this section we will introduce the essential theoretical ideas [7, 8] on which the numerical method is based. For the sake of simplicity, consider the compact Abelian model with staggered fermions which, in four dimensions, describes four flavors (the generalization to other gauge models as well as Wilson fermions is straightforward). We will mention briefly the differences with the noncompact Abelian model that we have also studied extensively in [9]. The action for the compact case is

$$\begin{aligned} S = & \tfrac{1}{2}\sum_{x,\mu}\eta_\mu(x)\bar{\chi}(x)\{U_\mu(x)\chi(x+\mu) \\ & -U^*_\mu(x-\mu)\chi(x-\mu)\} \\ & +m\sum_x\bar{\chi}(x)\chi(x)-\beta\sum_{x,\mu<\nu}\mathrm{Re}U_{\mu\nu}(x), \end{aligned} \tag{1}$$

where

$$U_{\mu\nu}(x)=U_\mu(x)U_\nu(x+\hat{\mu})U^*_\mu(x+\hat{\nu})U^*_\nu(x). \tag{2}$$

In the above, $\beta=1/e^2$ and the Kogut-Susskind fermion fields χ and $\bar{\chi}$ are coupled to the fields A_μ through the compact link variable $U_\mu(x)=e^{ieA_\mu(x)}$; $\eta_\mu(x)$ are the usual Kogut-Susskind phases. The bare fermion mass m is expressed in lattice units. The corresponding partition function, for n_f flavors, is

$$\begin{aligned} \mathcal{Z} &= \int[d\chi][d\bar{\chi}][dU_\mu(x)]e^{-S} \\ &= \int[dU_\mu(x)]\,[\det\Delta(m,U_\mu(x))]^{\frac{n_f}{4}}\,e^{-S_G}, \end{aligned} \tag{3}$$

where $\det\Delta(m,U_\mu(x))$ is the determinant of the fermion matrix, obtained after integrating out the Grassman fermion fields $\bar{\chi}$ and χ. It is a gauge-invariant object with a complicated dependence on gauge field operators, such as the plaquette energy, larger closed Wilson loops, and Polyakov lines. S_G is the pure gauge action.

We first define the density of states at fixed pure gauge energy as

$$N(E)=\int[dU_\mu(x)]\delta\left(\sum_{x,\mu<\nu}\mathrm{Re}U_{\mu\nu}(x)-6VE\right). \tag{4}$$

In compact models the above expression is well defined on a finite lattice; on the contrary, if the gauge variables have a noncompact support it is divergent even on a finite lattice, due to the divergence of the gauge group integration. This problem can be overcome either by gauge fixing or by factorizing the divergence, as has been done in [9].

In terms of $N(E)$, the partition function can be written as a one-dimensional integral

$$\mathcal{Z}=\int dE N(E)e^{6\beta VE}e^{-S^F_{\mathrm{eff}}(E,m)}, \tag{5}$$

where the effective fermionic action $S^F_{\mathrm{eff}}(E,m)$ is given by the expression

$$\begin{aligned} e^{-S^F_{\mathrm{eff}}(E,m)} &= \langle[\det\Delta(m,U_\mu(x))]^{\frac{n_f}{4}}\rangle_E \\ &= \frac{\int[dU_\mu(x)][\det\Delta(m,U_\mu(x))]^{\frac{n_f}{4}}\delta\left(\sum\mathrm{Re}U_{\mu\nu}(x)-6VE\right)}{\int[dU_\mu(x)]\delta\left(\sum\mathrm{Re}U_{\mu\nu}(x)-6VE\right)}; \end{aligned} \tag{6}$$

i.e., it is related to the logarithm of the value of the termionic determinant averaged over gauge configurations at fixed pure gauge (plaquette) energy E. The total effective action is therefore

$$S_{\mathrm{eff}}(E,V,\beta,m)=-\ln N(E)-6\beta VE+S^F_{\mathrm{eff}}(E,m), \tag{7}$$

where we have included the contribution from the density of the states in the effective action. Note that, if the fermionic effective action is linear in E, the effect of the fermions merely amounts to a shift in β. We shall also be using the effective pure gauge action defined by

$$S^G_{\mathrm{eff}}(E,V,\beta,m)=-\ln N(E)-6\beta VE. \tag{8}$$

There are two ways in which these analytical manipulations can be implemented in practical simulations. In both ways, $S^F_{\rm eff}(E,m)$ is obtained numerically for a number of E values. This is done by generating configurations at fixed energy E, according to (6) (see below for details). The value of $S^F_{\rm eff}$ at intermediate E values is obtained by the use of a standard interpolating routine.

The first possibility is to exploit the fact that the full effective action depends only on the plaquette energy E, in order to perform a change of integration variables in (5), going back from E to the original link variables $U_\mu(x)$. Then canonical simulations of the corresponding effective theory are possible (see Refs. [7, 8] for details). Almost all our compact U(1) simulations were carried out using this variant of our method [7, 8, 11].

The second possibility is to perform separate numerical evaluations of $N(E)$ and $S^F_{\rm eff}(E,m)$ as functions of E and obtain the partition function (5) as a one-dimensional integral. In the case of the noncompact Abelian theory, for which the underlying gauge theory is quadratic, the density of states is known analytically [9].

One advantage of these procedures is that the effective fermionic action (6) does not depend on the inverse coupling constant β and therefore we do not need to repeat the fermionic simulations for each different value of β. Since the dependence on n_f is trivial, we also do not need to repeat the simulation for each n_f value. The second advantage is that it is now possible to tackle the problem of statistical correlations between measurements of the fermionic determinant. This is because when simulating Eq. (6) for the calculation of $S^F_{\rm eff}(E,m)$, we need to generate pure gauge field configurations at fixed energy. This is done by microcanonical methods (see Sec. VIII for details). The cost of their implementation is negligible when compared to that of the computation of the fermionic determinant. This enables us to perform a huge number of microcanonical sweeps between measurements of the determinant, thus decorrelating in a satisfactory manner.

III. FEASIBILITY OF A DIRECT COMPUTATION OF THE EFFECTIVE ACTION

The first step in the practical application of our method consists in the determination of the effective fermionic action (6), namely, in the evaluation of the average of the fermionic determinant over configurations at fixed pure gauge energy. As has been outlined above, this is done by generating these gauge configurations with a flat probability distribution function (PDF). The problem is that the fermionic determinant (i.e., the operator that we want to measure with this PDF) varies widely over its range of values and is not an extensive quantity (i.e., it diverges exponentially with the lattice volume). Therefore, we consider the logarithm of the fermionic determinant, which is an extensive quantity, diverging linearly with the lattice volume V. On general grounds, we expect it to have a PDF which may easily be sampled numerically. In fact, at fixed pure gauge energy E, the PDF of $\ln\det\Delta$ is numerically found to be bell shaped, with a sharp peak.

To be definite, consider the logarithm of the fermionic determinant per unit volume y with a PDF $p_E(y,V)$ at fixed pure gauge energy E, mass m, and lattice volume V. The average value of the fermionic determinant is then given by

$$\langle\det\Delta\rangle_E = \int p_E(y,V)e^{Vy}dy. \tag{9}$$

In the thermodynamic limit, the dominant contribution to the effective fermionic action comes from the integration region of (9) around the location y_1 of the maximum of the integrand. The position of this maximum is determined by the occurrence of two competing effects.

The probability distribution of the logarithm of the fermionic determinant reaches a maximum at y_0 and decays from this value with some law.

The second factor in the integrand grows exponentially with y.

In practical simulations, where the statistics is obviously limited, the probability distribution $p_E(y,V)$ will be sampled with reasonable precision for values of y near y_0. The distribution far from the maximum, on the other hand, is difficult to evaluate numerically.

We will now investigate which conditions must be satisfied by a numerical simulation in order to make possible a reliable computation of the mean value of the determinant and its relative error. To this end, we will make the hypothesis that the functional form of the PDF of the logarithm of the determinant $p_E(y,V)$ is known.

Let $\{x_1,\ldots,x_N\}$ be an ensemble of independent measurements of the fermionic determinant x. The central limit theorem does guarantee that the logarithm of the average of x,

$$B = \ln\left(\frac{1}{N}\sum_{i=1}^{N}x_i\right), \tag{10}$$

goes to $\ln\langle x\rangle$ with an error inversely proportional to the square root of N, for N large enough. However, it may happen that the minimum value of N for which this is true is so large that the measurement is not feasible.

A direct analytic calculation of the dispersion of B is very complex even for a simple $p_E(y,V)$, but a good estimator of the error can be

$$\Delta B \equiv \ln\left(\langle x\rangle + \frac{\sigma_x}{\sqrt{N}}\right) - \ln\langle x\rangle = \ln\left(1+\frac{\sigma_x}{\langle x\rangle\sqrt{N}}\right). \tag{11}$$

In the $N\to\infty$ limit ΔB behaves as expected

$$\frac{\sigma_x}{\langle x\rangle\sqrt{N}} \ll 1 \longrightarrow \Delta B \approx \frac{\sigma_x}{\langle x\rangle\sqrt{N}}. \tag{12}$$

But when N is small, the behavior can be radically different:

$$\frac{\sigma_x}{\langle x\rangle\sqrt{N}} \gg 1 \longrightarrow \Delta B \approx \ln\frac{\sigma_x}{\langle x\rangle} - \frac{1}{2}\ln N; \tag{13}$$

i.e., the error is almost constant in N, decreasing as $\ln\sqrt{N}$. The change in the behavior takes place approximately at

$$N_0 \equiv \frac{\sigma_x^2}{\langle x \rangle^2}. \tag{14}$$

It is interesting to evaluate the value of N_0 for some simple PDF's and to compute its numerical value using the results of real simulations. Let us first assume that the PDF of the logarithm of the fermionic determinant is a Gaussian of the form

$$p_E(y,V) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left[-(y-y_0)^2/(2\sigma^2)\right]. \tag{15}$$

In this case

$$\ln\langle \det \Delta \rangle_E = y_0 + \frac{\sigma^2}{2} \tag{16}$$

and

$$N_0 = e^{\sigma^2} - 1. \tag{17}$$

In the case of a Poisson distribution of the form

$$p_E(y,V) = \frac{\lambda^{n+1}}{n!} y^n e^{-\lambda y}, \tag{18}$$

we have

$$N_0 = \left[1 + 1/(\lambda^2 - 2\lambda)\right]^{n+1} - 1. \tag{19}$$

To have an idea of the numbers involved, we have fitted with a Gaussian (Poisson) distribution the PDF of the logarithm of the fermionic determinant obtained from a numerical simulation in compact U(1) at lattice size $L = 8$, energy $E = 0.5$, and bare fermion masses $m = 0.0$, 0.1, and 0.2. The results are

$$\begin{aligned} m = 0 &\longrightarrow N_0^{\text{Gauss}} = 10^{12}, \quad N_0^{\text{Poisson}} = 10^{16}, \\ m = 0.1 &\longrightarrow N_0^{\text{Gauss}} = 10^5, \quad N_0^{\text{Poisson}} = 2 \times 10^5, \\ m = 0.2 &\longrightarrow N_0^{\text{Gauss}} = 600, \quad N_0^{\text{Poisson}} = 700, \end{aligned} \tag{20}$$

which, except for very large masses, show the nonfeasibility of the direct computation, since in practice it is difficult to have N larger than a thousand. Moreover, the situation worsens rapidly with increasing V.

In conclusion, we have seen that the extremely slow increase of the statistical significance of measurements of the determinant (9) is related to the fact that the PDF of its logarithm is weighted with an exponential function; in this way its determination depends on the behavior of the tails of the PDF.

IV. EFFECTIVE ACTION AND THE n_f EXPANSION

In order to tackle the problems we have exposed above, we propose a less ambitious procedure which consists in the evaluation of the first few moments of the distribution. This can be done by expanding the effective action in powers of the flavor number n_f [9], and by computing in practice only the first contributions to this expansion. In particular, the effective action (6) is related to the average of the fermionic determinant, computed over gauge field configurations at fixed pure gauge energy. In general, for n_f species, the effective fermionic action

$$-S_{\text{eff}}^F(E,m,n_f) = \ln\langle e^{\frac{n_f}{4} \ln\det\Delta(m,U_\mu(x))} \rangle_E \tag{21}$$

can be expanded in cumulants as

$$\begin{aligned} -S_{\text{eff}}^F(E,m,n_f) = {} & \frac{n_f}{4}\langle \ln\det\Delta(m,U_\mu(x)) \rangle_E \\ & + \frac{n_f^2}{32}\{\langle (\ln\det\Delta)^2 \rangle_E - \langle \ln\det\Delta \rangle_E^2\} \\ & + \frac{n_f^3}{384}\{\langle (\ln\det\Delta - \langle \ln\det\Delta \rangle_E)^3 \rangle_E\} \\ & + \cdots, \end{aligned} \tag{22}$$

which is nothing but an expansion in powers of the flavor number of the effective fermionic action. Successive terms in the expansion represent higher-order fluctuations, and are, in general, increasingly difficult to evaluate. In shorthand notation we can write the above as

$$\ln\langle (\det\Delta)^{\frac{n_f}{4}} \rangle_E = \sum_{i=1}^{\infty} \left(\frac{n_f}{4}\right)^i \frac{k_i}{i!}, \tag{23}$$

where k_i is the ith cumulant of the distribution $p_E(y,V)$. Note that up to $i = 1$ the series corresponds to the approximation $\ln\langle \det\Delta \rangle \approx \langle \ln\det\Delta \rangle$, while for a Gaussian distribution the cumulants are zero for $i > 2$. In general, the cumulant expansion has infinite terms: For instance, for a Poisson distribution $k_i = (i-1)!\frac{n+1}{\lambda^i}$.

In the simulations reported below, only a few cumulants are evaluated. The truncation of the expansion thus introduces systematic errors, in addition to statistical ones.

Figures 1 and 2 show the results for the compact U(1) model, with $L = 8$ and $m = 0.0$, 0.1, respectively. We plot the logarithm of the averaged determinant ($n_f = 4$) for various orders of the cumulant expansion. We also show results obtained with a Poisson fit of the PDF of the logarithm of the fermionic determinant. In all the cases analyzed (lattices up to 10^4 and various values of

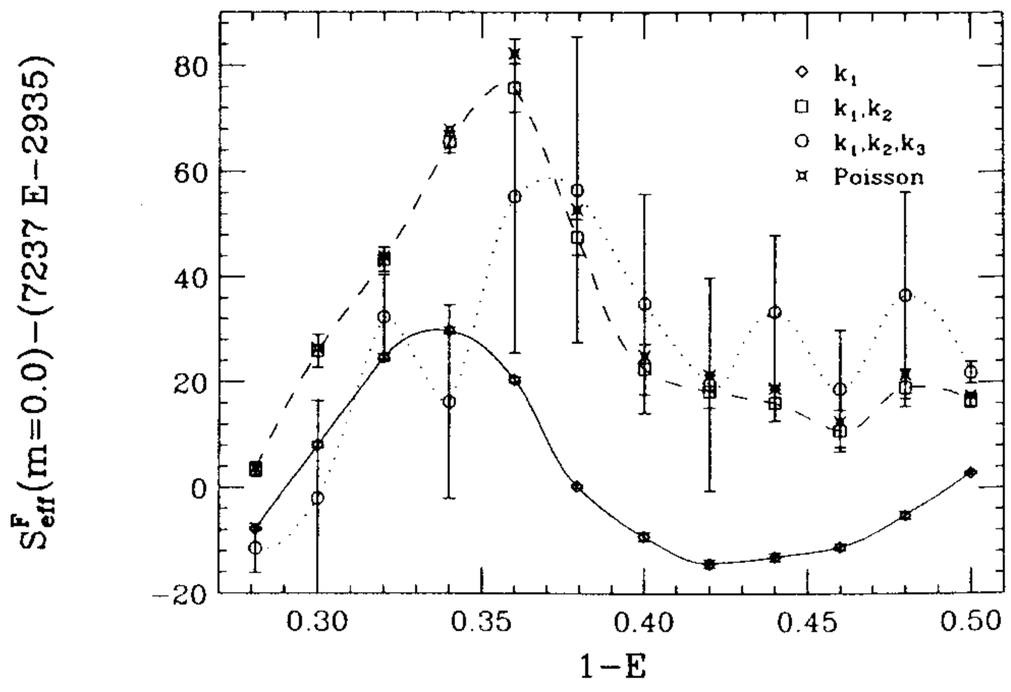

FIG. 1. Logarithm of the average determinant in an $L = 8$ lattice, with $n_f = 4$, $m = 0$ at the first ($k1$), second ($k1, k2$), and third order of the cumulant expansion, as well as for a Poisson distribution. In this last case error bars are not shown for clarity; they are very similar to the second order ones. The y axis is the action with a linear shift (of the results of the first order) to increase visibility.

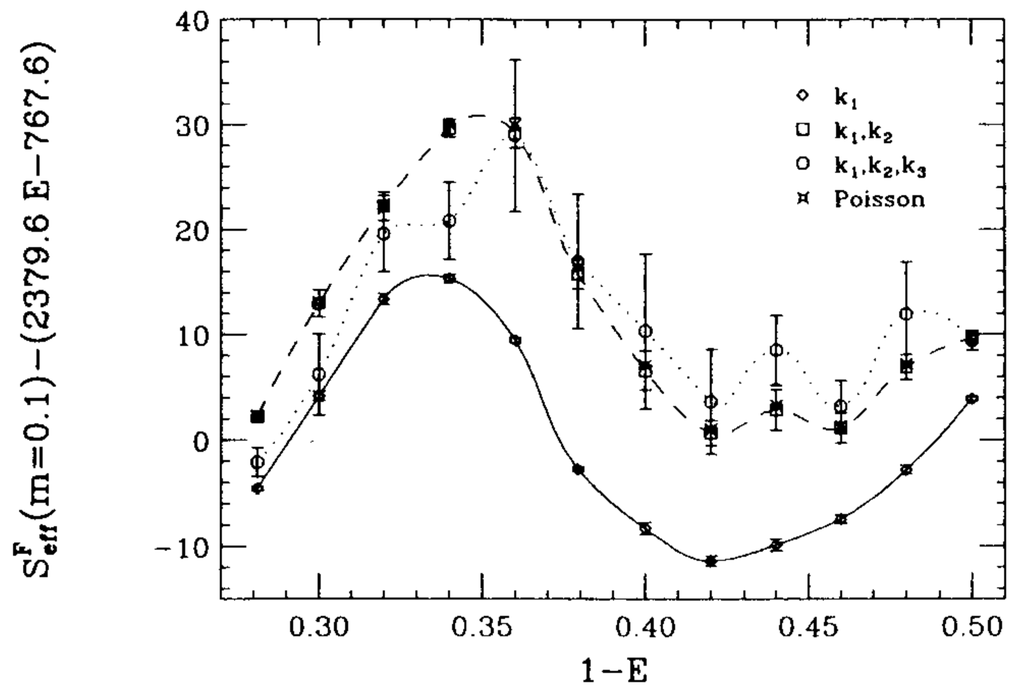

FIG. 2. Same as Fig. 1, but at $m = 0.1$

m), an apparent stability settles in after the second order.

The relevance of systematic errors will now be discussed for the special case of noncompact QED [9] where we have the best statistics. Figure 3 is a plot of the PDF of the logarithm of the fermionic determinant in a 8^4 lattice, $m = 0.0$, and normalized pure gauge energy $E = 1.20$ (note that $0 \leq E \leq \infty$ in the noncompact case, $-1 \leq E \leq 1$ in the compact one). This case has been chosen because the statistics are particularly good (1300 configurations). We stress that every measurement (i.e., every diagonalization of the fermion matrix needed for the calculation of the determinant) is separated from the previous one by 1000 iterations of a canonical Monte Carlo (MC) process, followed by an appropriate rescaling of the gauge fields in order to bring the energy to the required value. This procedure guarantees the decorrelation of the successive gauge configurations on which measurements of the determinant are performed. Coming back to Fig. 3, the solid line is a Gaussian fit of the distribution measured numerically. For the range of masses we are interested in ($0 \leq m \leq 0.1$), the goodness of the fit is evident and independent of m. If, from these results, we assume that the probability distribution of the logarithm of the fermionic determinant at fixed pure gauge energy is Gaussian, then only the first two contributions to the effective fermionic action will be different from zero and no systematic errors will be introduced by the truncation of the expansion.

In Figs. 4(a)–4(c) we present our results for the first three contributions to $S^F_{\rm eff}$, respectively, as a function of energy at $m = 0$. The results for the third contribu-

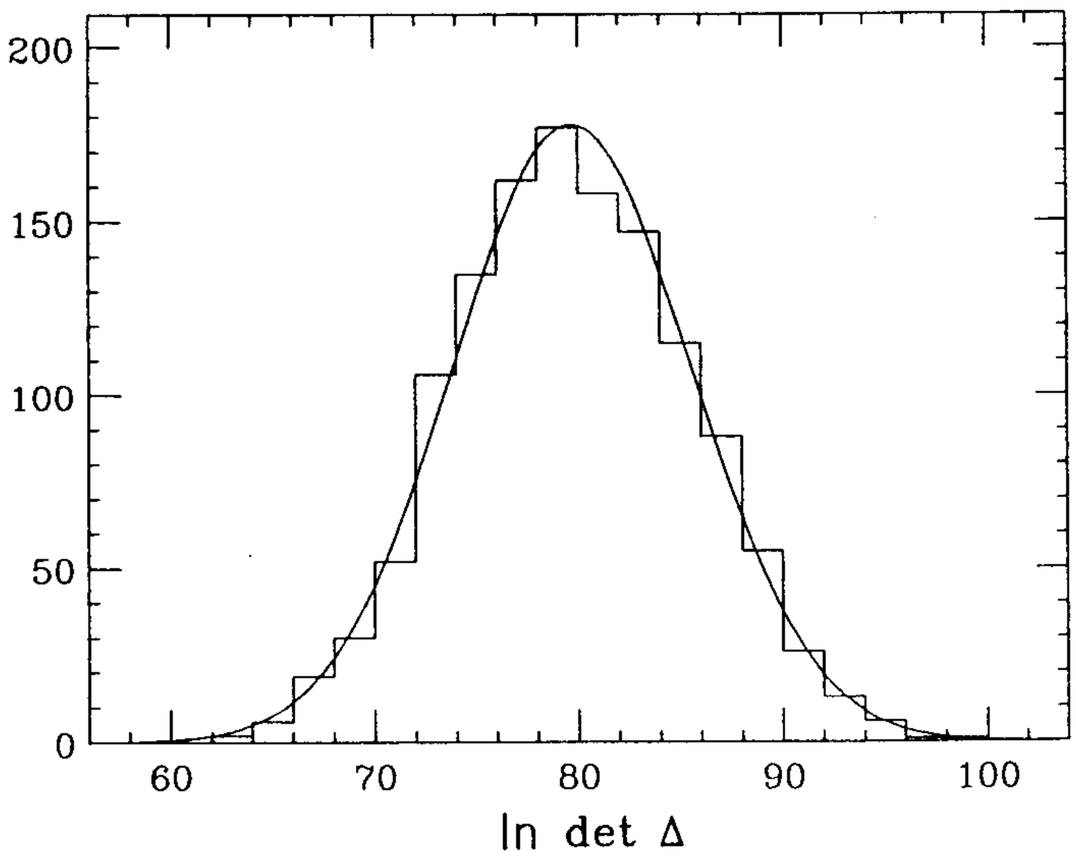

FIG. 3. Probability distribution of $\ln \Delta(m, A_\mu)$ at $m = .0$, $E = 1.20$, and $N_f = 4$ for an 8^4 lattice. The solid line is Gaussian fit, with $\chi^2/N_{\rm DF} = 0.487$.

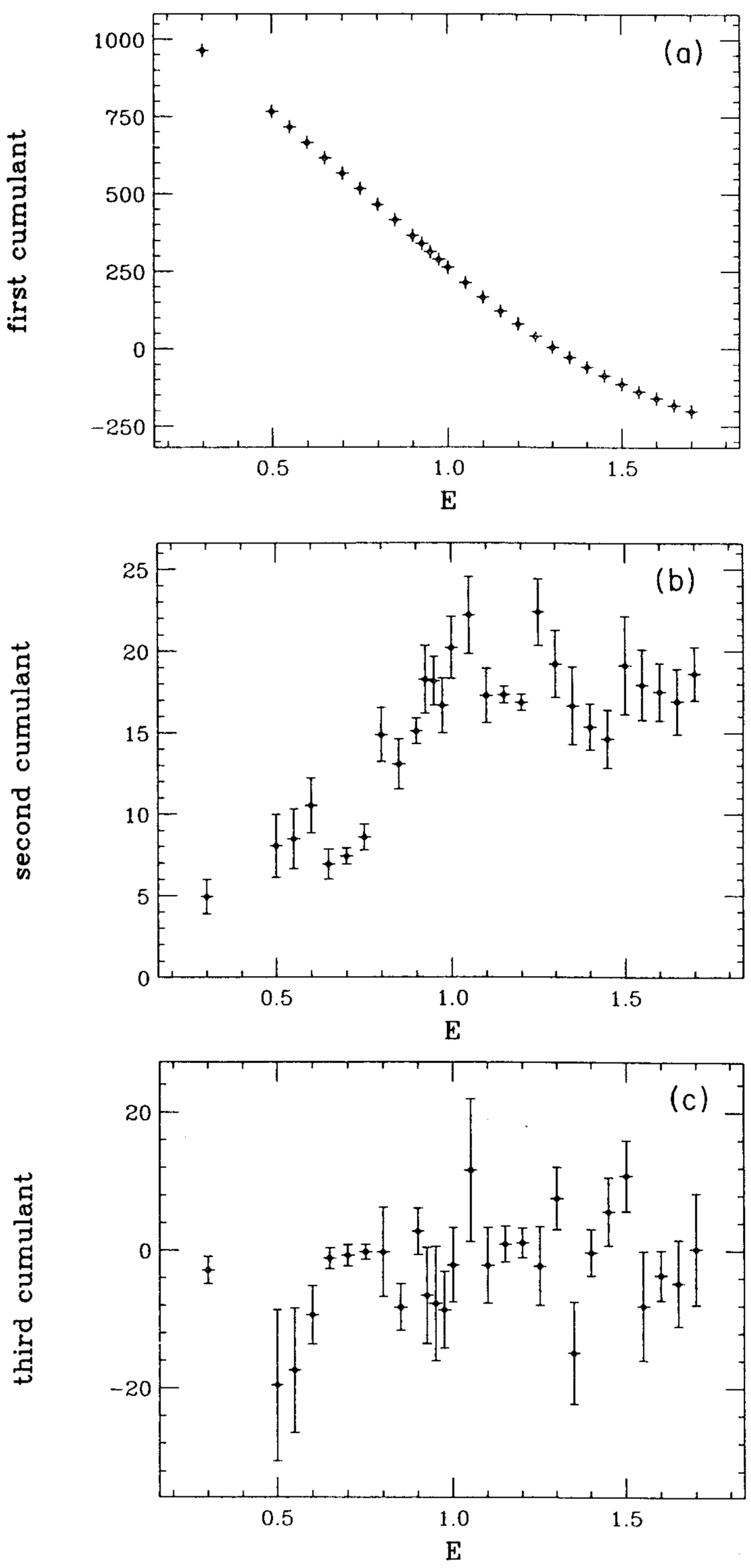

FIG. 4. (a) First, (b) second, and (c) third cumulant of $S^{\rm eff}_f$ vs E at $m = 0$ for an 8^4 lattice.

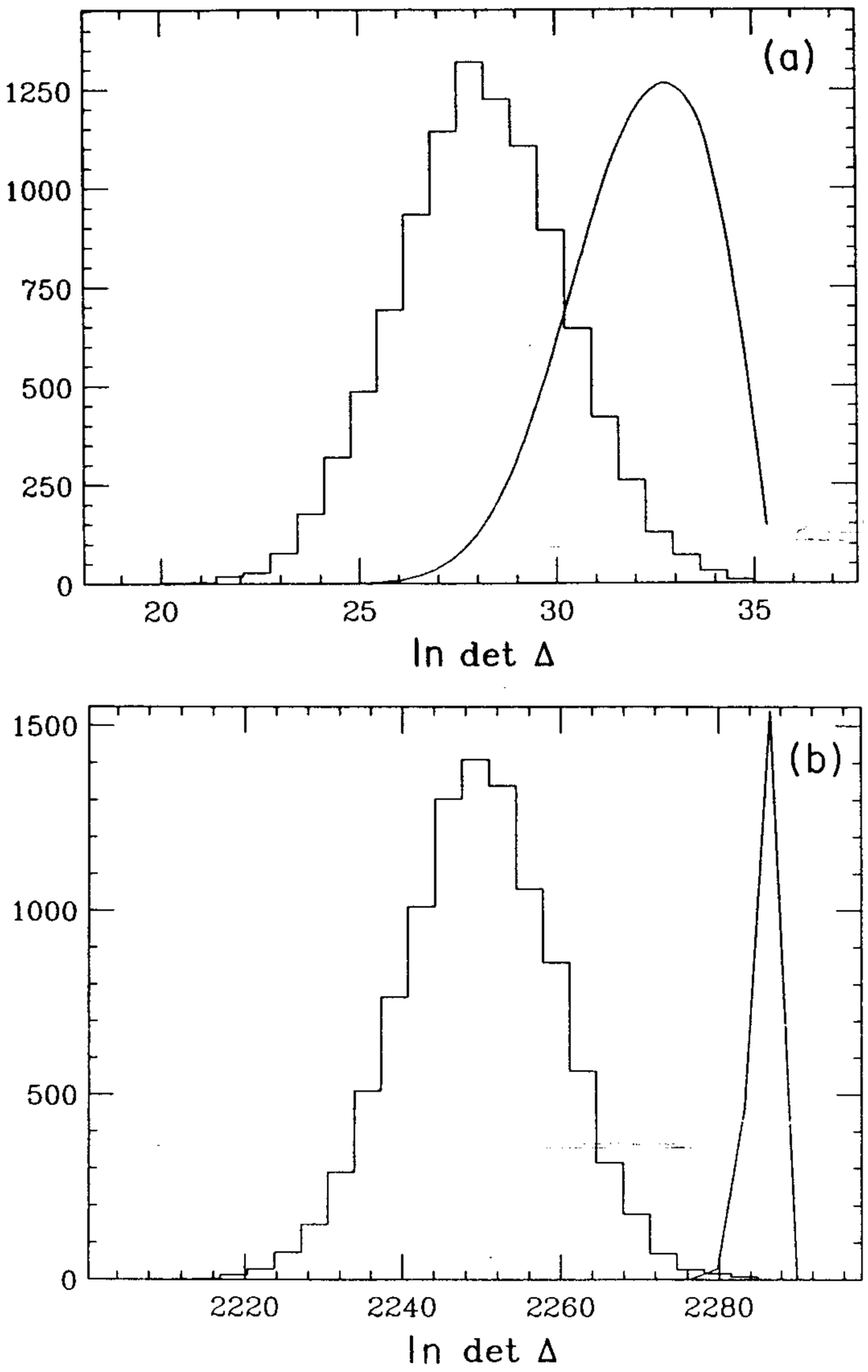

FIG. 5. PDF of the logarithm of the determinant and its product with the exponential in (9) for $V = 4^4$ (a) and $V = 12^4$ (b).

tion are compatible with zero, according to the previous discussion.

At this point it is important to comment on the dependence of the various terms of the expansion on the lattice volume, since at the end one is interested in the thermodynamic limit $V \to \infty$. Since the logarithm of the determinant is bounded, it is easy to show that $\ln\langle \det \Delta \rangle_E \sim V$. The same argument suggests that $\langle \ln \det \Delta \rangle_E \sim V$, and that successive terms in the expansion diverge, *at most* as V. Hence a situation is conceivable in which the analysis of the previous paragraph would imply that, although the direct evaluation of the average determinant is impossible, the first term in the cumulant expansion gives the exact result in the thermodynamical limit.

To illustrate this possibility, we have taken the data of the compact model in the 4^4 lattice, at $m = 0.1$ and $E = 0.5103$, fitted the PDF with a Gaussian, and extrapolated this fit to larger volumes assuming that $\sigma^2 \sim V^{0.7}$. In Figs. 5(a) and 5(b) we present the PDF of the logarithm of the determinant and its product with the exponential in (9) for $V = 4^4$ and 12^4, respectively. It is evident that the direct evaluation of the average determinant in the last case would fail, while the first term of the expansion (22) would give the exact result in the thermodynamic limit.

V. n_f EXPANSION FOR PHYSICAL OBSERVABLES

It must be stressed that the effective action does not have a direct physical meaning, and so the fact that systematic errors in its evaluation are small does not guarantee *a priori* the correctness of numerical determination of other physical quantities. The possibility of evaluating physical quantities by the method used for the effective action can be studied by means of power expansions similar to (22). To be more specific, let us consider the case of the chiral condensate in the Abelian model [9].

The vacuum expectation value $\langle \bar{\psi}\psi \rangle$ can be computed as

$$\langle \bar{\psi}\psi \rangle = -\frac{1}{V} \frac{\int dE e^{-S_{\rm eff}(E,m,\beta,n_f,V)} \frac{\partial}{\partial m} S^F_{\rm eff}(E,m,n_f)}{\int dE e^{-S_{\rm eff}}}; \tag{24}$$

namely, the chiral condensate is the average value of the derivative with respect to the mass of the normalized fermionic effective action, with a probability distribution deriving from the total effective action.

The expansion of the effective fermionic action in powers of the number of flavors (22) leads to a similar expansion for the contributions to the chiral condensate:

$$-\frac{\partial S^F_{\rm eff}}{\partial m} = \frac{n_f}{4}\langle {\rm Tr}\Delta^{-1}\rangle_E + \frac{n_f^2}{16}\langle (\ln\det\Delta - \langle \ln\det\Delta\rangle_E)({\rm Tr}\Delta^{-1} - \langle {\rm Tr}\Delta^{-1}\rangle_E)\rangle_E + \frac{n_f^3}{64}\langle (\ln\det\Delta - \langle \ln\det\Delta\rangle_E)^2({\rm Tr}\Delta^{-1} - \langle {\rm Tr}\Delta^{-1}\rangle_E)\rangle_E + \cdots. \tag{25}$$

Therefore the chiral condensate is given by the average value over the probability distribution implicit in (24) of the successive terms in (25), normalized by V.

Here, as in the computation of the effective fermionic action, the difficulty with which successive terms in the expansion (25) can be numerically evaluated, increases with the expansion order. In practice, also in this case we will be forced to truncate the expansion to a certain order. Thus the evaluation of the chiral condensate will also be affected in principle by systematic errors due to

this approximation.

However, following the previous analysis for the effective fermionic action, the only nonzero contributions to the chiral condensate are the first two in (25), assuming that the PDF of the logarithm of the determinant at fixed pure gauge energy is Gaussian. The results obtained from the first two contributions in (25) for the noncompact lattice QED reported in [9] were in very good agreement with results obtained by other groups using the hybrid MC technique [18]. Thus, the truncation of the series appears to be justified.

To obtain expressions similar to (25) for any other operator O, which is not a term of the action, it is enough to introduce an external source coupled to O in the original action. Differentiation gives

$$\langle O \rangle = \frac{\int dE e^{-S_{\rm eff}(E,m,\beta,n_f,V)} O_E}{\int dE e^{-S_{\rm eff}}}, \tag{26}$$

with

$$O_E = \langle O \rangle_E + \frac{n_f}{4} \langle (\ln \det \Delta - \langle \ln \det \Delta \rangle_E)(O - \langle O \rangle_E) \rangle_E + \frac{n_f^2}{16} \langle (\ln \det \Delta - \langle \ln \det \Delta \rangle_E)^2 (O - \langle O \rangle_E) \rangle_E + \cdots . \tag{27}$$

In this way the average value of any operator can be obtained as a power series in n_f. The convergence of this type of expansion will be better at smaller n_f, and/or for a narrower PDF $p_E(y,V)$ of the logarithm of the fermionic determinant. This point will be further considered in Sec. VII.

VI. SOME EXAMPLES OF PHYSICAL APPLICATIONS

In this section we briefly review some results which have already been obtained through the use of our method [7–9]; we also present some unpublished ones. Let us first describe the general realization of simulations.

As described above, we first produce configurations at fixed pure gauge energy through a microcanonical process. Next, the fermionic matrix is exactly diagonalized at zero mass, by finding all its eigenvalues, using a modified Lanczos algorithm [19]. This allows us to compute once and for all the fermionic effective action at all mass values, including $m = 0.0$ (see Sec. VIII for details). This part of the computation is the most costly in terms of CPU and the time needed is of order V^2. It is conceivable that other faster methods could be used to compute the fermionic determinant at fixed mass, especially at large volumes. In this case the fermionic calculation should be repeated for each value of the mass. From the eigenvalues we compute the fermionic effective action for several fixed values of the pure gauge energy. We subsequently use numerical interpolation of these results in a canonical simulation of the effective theory.

Alternatively, in the noncompact case, since the density of states is known analytically, a one-dimensional numerical integral suffices to extract physical quantities. In all the cases considered we have checked that the choice of interpolation has no effect on the physical results.

In [7, 8] we studied in detail the compact U(1) model. In this model the pure gauge theory shows a first-order phase transition at finite coupling [20–22]. We refer the interested reader to Refs. [7, 8] for more details on the results obtained in the unquenched case. Here it suffices to say that our results agree extremely well with results obtained with standard methods [17]. Entirely new are the results at $m = 0$, which cannot be obtained by other methods. In particular, it was shown that the strength of the transition increases at decreasing mass: At $m = 0$ the system managed to make a single tunneling in 200 000 iterations [8].

As an example we analyze here the shift of the value of the critical coupling of the deconfining transition in compact U(1) and the latent heat. Since the tunneling between the two states at the first-order phase transition point is very difficult at $m = 0.0$, it is preferable to do this kind of analysis directly on the effective action (7). In this way we will also have a better understanding of the reason for the strengthening of the transition.

The technical problem in this case is that for large lattices a simulation at fixed β, needed to evaluate numerically $N(E)$, generates configurations with energies in an exceedingly narrow interval. We have solved this problem by performing simulations at different β in such a way that the energy intervals are superposed. In this way it is possible to use the method proposed in [23] for extracting $N(E)$. Note that in the case of noncompact QED the situation is much more favorable since the effective pure gauge action can be computed analytically.

We computed the total effective action (7) for $L = 8$ ($n_f = 4$) and different masses. In Fig. 6 the effective action (7) at $m = 0$ in the second-order approximation

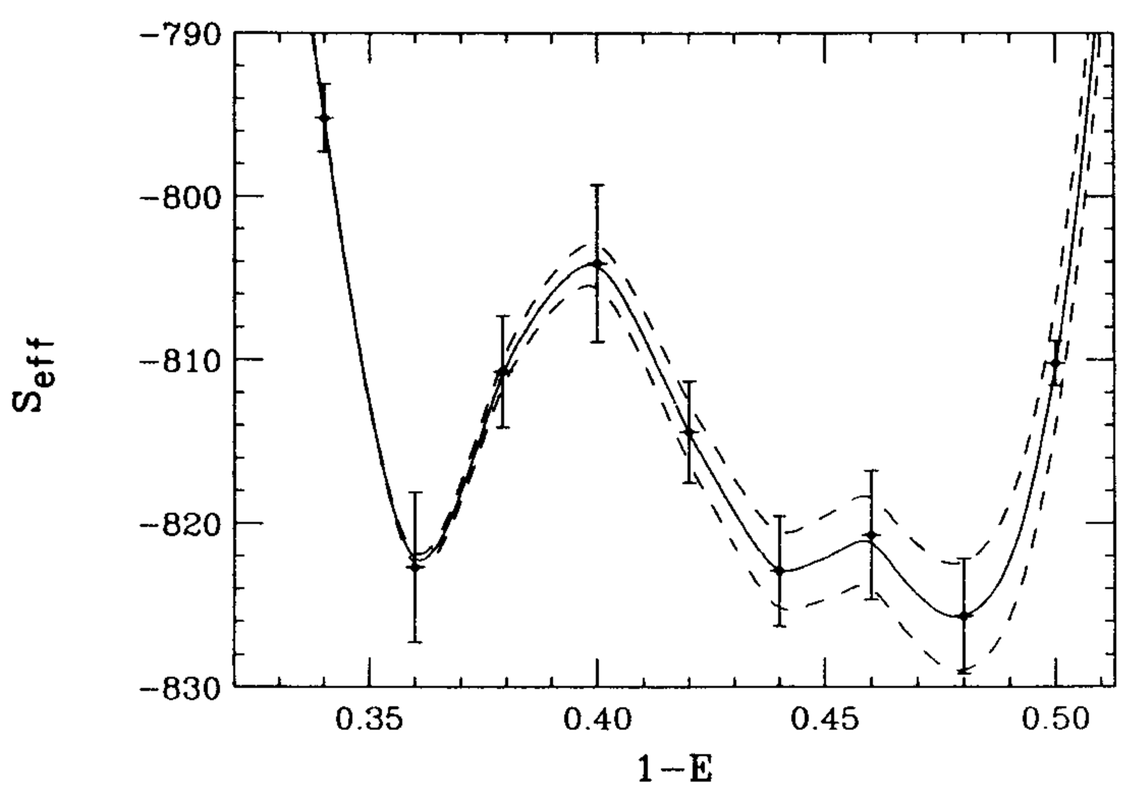

FIG. 6. Effective potential $S_{\rm eff}(E,V,\beta,m)$ for $L = 8$, $n_f = 4$, and $m = 0$ obtained by summing the pure gauge contribution computed by folding different Monte Carlo simulations, and the fermionic action in the Gaussian approximation (cumulants expansion up to k_2). The solid line corresponds to $\beta = 0.853$ (our determination of β_c). Dashed lines correspond, respectively, to $\beta = 0.852$ and $\beta = 0.854$ and are shown to allow an estimation of the error.

(Gaussian distribution) is reported.

In Fig. 7 we report the change of β_c from the pure gauge value, as a function of the order in the cumulant expansion for various masses. Again, one observes an apparent stability from the second-order contribution onwards.

In the case of the latent heat (Fig. 8) the errors are considerably larger, but the conclusions remain unchanged.

The results for the effective fermionic action allow us to understand why the pure gauge transition becomes stronger when dynamical fermion effects are included: In fact, the fermionic effective action shows a convexity for energies between the two minima of the pure gauge effective action.

For m large enough (i.e., $m = 0.1$), where tunneling between the states is easier, it is still possible to extract the same observable by means of a standard canonical simulation, including the effective fermionic action in the Metropolis algorithm in a straightforward way. This possibility can be used as a consistency check of the results shown above.

We used in these canonical simulations four different forms of the effective fermionic action.

(1) Average logarithm of the fermionic determinant (first term in the cumulant expansion).

(2) Average logarithm and first fluctuation (first and second cumulants, i.e., Gaussian approximation).

(3) Same as case (2), but smoothing in energy the second cumulant. This procedure is performed to avoid (nonphysical) fluctuations of the first derivative of the fermionic effective action S^F_{eff}, due to the use of different ensembles of configurations as samples for very near values of the pure gauge energy.

(4) Average logarithm plus the smoothing of the sum of the other cumulants. This is done in a spirit similar to that of case (3).

For these four choices of the action we have performed extensive canonical simulations (typically 70000 cycles,

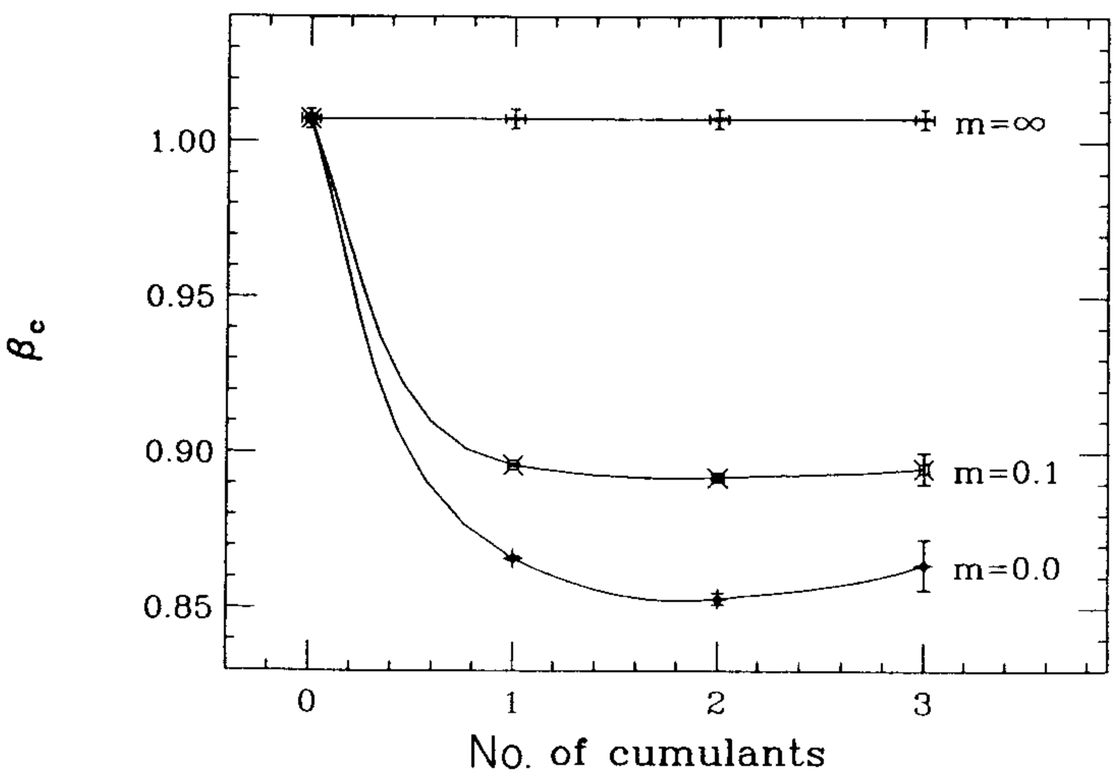

FIG. 7. Measurement of β_c in an $L = 8$ lattice for compact U(1) with $n_f = 4$. The x axis corresponds to the order of the approximation in the cumulant expansion. The values corresponding to the quenched ($m = \infty$), $m = 0.1$, and $m = 0.0$ cases are presented.

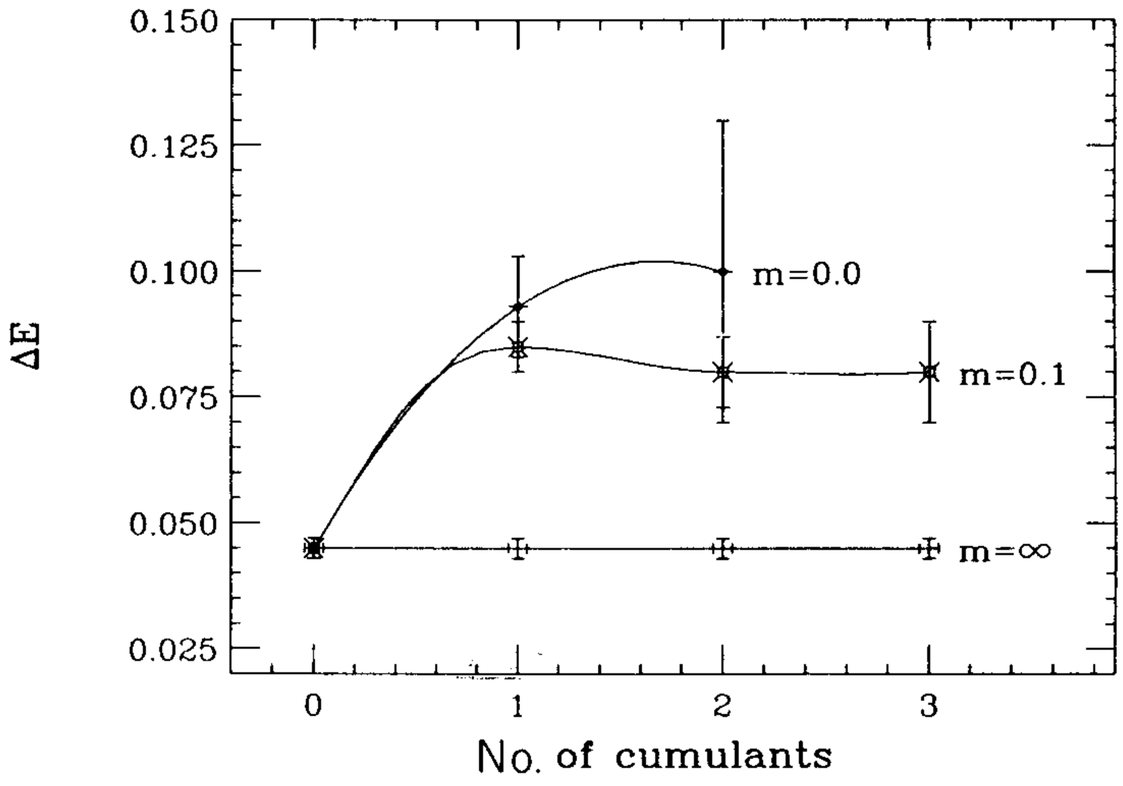

FIG. 8. Latent heat as a function of m.

each consisting in a three-hit Metropolis plus four overrelaxation sweeps [8]) at $\beta_0 = 0.8935$, a value near enough to the estimated critical coupling to allow the use of the Ferrenberg-Swendsen [23] procedure to reconstruct thermodynamical quantities in the whole critical region.

The main results are reported in Table I. As can be seen, we always have consistency with the results obtained with the effective potential method.

Finally, the method described in the previous paragraphs, including the expansion in cumulants, has been used in [9] for an extensive analysis of the noncompact Abelian model. This model is particularly interesting: While the pure gauge theory is quadratic and, hence, trivial, when dynamical fermions are added it develops a second-order transition at finite coupling [24], where the continuum limit might reside, possibly giving an example of a nonasymptotically free theory [9, 25, 26]. As for the method we propose, the particularity here is that the density of states $N(E)$ can be computed analytically, so that the computation of the partition function and its derivatives can be done through a one-dimensional numerical integration, and there is no need to perform a canonical simulation.

Again, our results on physical observables (plaquette energy, chiral condensate) agree very well with the ones existing in the literature. It is particularly important to note that, being able to obtain the plaquette energy at

TABLE I. Numerical results for the critical coupling β_c, height of the maximum of the specific heat h_c, Binder parameter B, and latent heat L_h. Each row is labeled in correspondence with the choice of one of the four fermionic effective actions enumerated in the text.

	β_c	h_c	B	L_h
1	0.893(2)	21(5)	0.663(1)	0.07
2	0.897(3)	27(7)	0.663(1)	0.07
3	0.893(2)	16(3)	0.664(1)	0.05
4	0.895(3)	16(3)	0.664(1)	0.06

$m = 0$, we can obtain the position of the phase transition in a way completely independent from the extrapolation of the chiral condensate at zero mass. The values of the critical coupling we obtain [$\beta_c = 0.224(3), 0.208(4)$ at $n_f = 2, 4$, respectively] agree with those reported in Refs. [25, 26].

VII. ISING AND COMPACT-NONCOMPACT QED MODELS

In the previous paragraphs we have introduced the method we propose, studied the limitations and approximations involved, and discussed some applications. This paragraph is more speculative; it attempts to explain, at least partially, the reasons for which the procedures described so far give reliable physical results.

From the discussion developed in the previous paragraphs, we can argue that a numerical determination of the expectation value of the fermionic determinant at fixed pure gauge energy is possible if the fluctuations of $\ln \det \Delta$ are small, i.e., if the PDF of this quantity at fixed energy has a very narrow width. This is in fact a condition for the fast convergence of the cumulant expansion (23). Taking into account just the first term in the cumulant expansion (23), we are neglecting the contribution to the effective fermionic action of all fluctuations, and this corresponds to some kind of mean field approximation. On the other hand, it is well known that mean field techniques give good results when the dimensionality of space-time d is large or the interaction is long ranged. In our case, $\ln \det \Delta$ is a nonlocal operator: We expect that its fluctuations at fixed energy are bounded.

A simple physical model (i.e., the Ising model) gives a very good illustration of these points. In this model, the partition function Z_N can be written as a sum over magnetizations M:

$$Z_N = \sum_M \sum_s \delta\left(\sum_i s_i - M\right) e^{-\beta H(s)} = \sum_M g_N(M) \langle e^{-\beta H(s)} \rangle_M, \tag{28}$$

where $g_N(M)$ indicates the number of configurations in a N spin lattice with a given magnetization M.

In this case the effective Hamiltonian $H_{\rm eff}(M)$ is a function of the magnetization M.

Let us consider two possible cases.

(i) The nonlocal Ising model:

$$H_{\rm nl} = \frac{-J}{N} \sum_{i \neq j} s_i s_j = \frac{J}{2} - \frac{J}{2N} M^2. \tag{29}$$

In this case the Hamiltonian depends only on the magnetization M, which implies that at fixed magnetization it does not fluctuate. This means that

$$\langle e^{-\beta H_{\rm nl}(s)} \rangle_M = e^{-\beta (H_{\rm nl}(s))_M}; \tag{30}$$

i.e., the only nonvanishing cumulant is the first one.

(ii) The local Ising model:

$$H_l = -J \sum_{\langle i,j \rangle} s_i s_j, \tag{31}$$

where now the sum is over nearest-neighbor spins. In this case the Hamiltonian, computed over configurations of fixed magnetization, suffers from large fluctuations. The effective Hamiltonian per unit volume can be defined through the relation

$$\langle e^{-\beta H_l(s)} \rangle_M = e^{N \bar{H}_{\rm eff}(M)} \tag{32}$$

N being the total number of spins.

A straightforward but tedious computation gives the first four cumulant contributions to the effective Hamiltonian per unit volume:

$$\begin{aligned} \bar{H}_{\rm eff}(M) &= dFm^2 + \frac{d}{2}F^2(1-m^2)^2 + \frac{d}{6}F^3(4m^2 - 8m^4 + 4m^6) \\ &\quad + \frac{d}{12}F^4[6d - 7 + (20 - 24d)m^2 - (36d - 10)m^4, \\ &\qquad -(12 + 24d)m^6 + (9 + 6d)m^8] \\ &\quad + \cdots, \end{aligned} \tag{33}$$

where d is the dimensionality of space, $m = M/N$, $F = J/KT$, and we have neglected contributions to the effective Hamiltonian which vanish in the thermodynamic limit.

Successive powers of F in the above expansion multiply increasingly high-order fluctuations of the Hamiltonian. Thus, from expression (33) it follows that in the local model the energy suffers from large fluctuations even at fixed magnetization, making physically relevant the contribution to the cumulant expansion of orders higher than the first.

In the nearest-neighbor Ising model, if we truncate the series (33) to lowest order, $N H_{\rm eff}(m^2) = \langle H_l(s) \rangle_M$, we get the mean field approximation. However, using the four terms in (33) gives much better physical results, especially for higher dimensions. In particular, in one and two dimensions we find a first-order phase transition, but our results in $d \geq 3$ approach those obtained in exact numerical simulations. In Table II we report the critical values of F and the order of the phase transition for several values of d. The results reported in this table show unambiguously how the convergence of the cumulant expansion improves significantly with increasing space di-

TABLE II. Critical values of F and order of the ferromagnetic phase transition for several values of the space dimensionality d. These values have been obtained using the four cumulant approximation to the effective Hamiltonian.

d	F_c	Order
2	0.42055	First
3	0.216076	Second
4	0.147672	Second
6	0.092023	Second
10	0.052836	Second

mensionality d. This is not surprising since, as stated above, the mean field approximation gives better results for increasing space dimensionality and/or long-ranged interactions, suggesting that the main criterion for the reliability of our numerical method in QFT might be the range of the interaction.

In the case of a quantum field theory (QFT) with dynamical fermions, the analogue of the Ising Hamiltonian is ln det Δ. Its nonlocal character may be responsible for the smallness of the fluctuations of this operator, as a function of the pure gauge energy, in accordance to what we observed in the numerical simulations.

The relevance of nonlocality for the convergence of the cumulant expansion can be further discussed in a simple, pure gauge Abelian lattice model with a mixed compact-noncompact action. In this model it is possible to perform an analysis similar to that for noncompact QED [9]. The effective action is therefore

$$e^{-S_{\rm eff}(E)} = \langle e^{-S_G(U_{x,\mu})} \rangle_E, \tag{34}$$

where $S_G(U_{x,\mu})$ is the compact contribution to the total action, and $\langle\rangle_E$ is now the average over configurations of fixed noncompact pure gauge energy E.

Figure 9 contains the numerical results for the first contributions to the cumulant expansion of (34) as a function of the noncompact energy E. As can be seen, the contribution of the second and third cumulants (or, more relevantly from a physical point of view, their derivatives with respect to E) are of the same order, or even larger than the first term. This is a completely different behavior from that of noncompact QED, in which the contribution of the second cumulant was less than 2%, the third cumulant being compatible with zero [9].

It is important to note that the compact gauge action which appears in the exponential of (34) possesses the full local symmetry of the noncompact action, as in QED. This has to be compared with the Higgs model at fixed norm of the scalar field:

$$e^{-S_{\rm eff}(E)} = \langle e^{-S_H(U_{x,\mu})} \rangle_E, \tag{35}$$

where $S_H(U_{x,\mu})$ is the Higgs term in the action. In this case the computation also fails, the failure being partly due to the fact that the Higgs part of the action, after integration of the scalar field, does not possess the local symmetry of the pure gauge part.

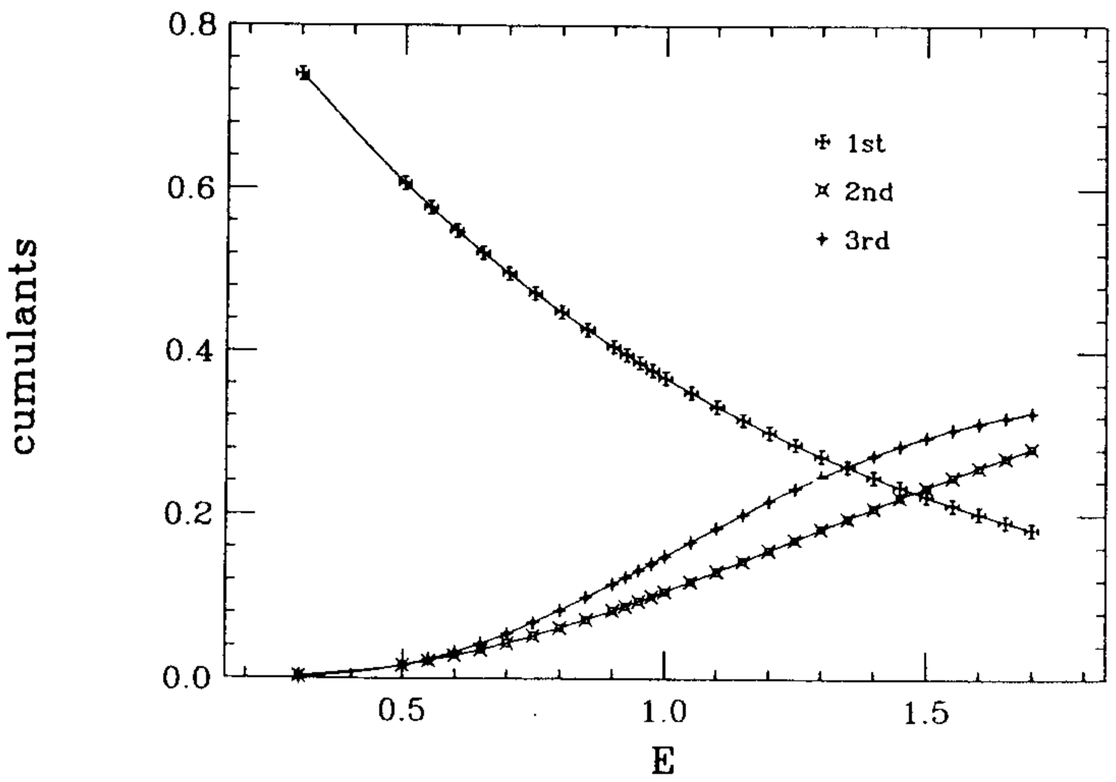

FIG. 9. First cumulants of $e^{-S_{\rm eff}(E)} = \langle e^{-S_G(U_{x,\mu})} \rangle_E$ [Eq. (34)].

The main difference between the model with a mixed action and the fermionic case is that the compact pure gauge action is local, whereas in QED the logarithm of the fermionic determinant, which appears in the expression analogous to (34) defining the fermionic effective action, is nonlocal in terms of the gauge variables.

These examples once more suggest that the nonlocal character of the operator which defines the effective action (effective Hamiltonian for the Ising model) is essential for the feasibility of the numerical computation of the fermionic effective action.

VIII. FERMIONIC NUMERICAL CALCULATIONS

As already stated, the realization in practical simulations is logically divided into the following parts: (a) generation of decorrelated configurations at fixed energy; (b) diagonalization of the fermionic matrix and determination of its zero mass eigenvalues; (c) canonical simulation of the effective theory.

We will now describe in some details the practical realization of the above points. As for point (a), the generation of configurations at fixed energy, i.e., the microcanonical process, can be realized in different ways, depending on the theory under consideration. In particular, for the compact U(1) case we have used the overrelaxation method [27] in which each link is changed deterministically, so that the energy remains fixed. This has a unique realization in U(1). In SU(2) the solution to the overrelaxation equations is not unique, and in general the choice of the new value for the link is done so that the distance from the previous value is maximized. For SU(3) overrelaxation does not exist, but microcanonical methods do [28]. Note that in any case, due to machine precision, after many iterations of any microcanonical procedure the energy will slightly change. To avoid this problem, after a fixed number of cycles some sweeps are performed, guided so that the energy is forced towards the correct value.

In the case of the noncompact Abelian model a different approach is used. In fact, the overrelaxation procedure would move the fields towards very large values since in this case they are not bounded. This will cause problems related to machine precision. In order to overcome the problem, we chose instead to exploit the fact that the pure gauge action is quadratic in the fields, while the measure is linear. Thus, a rescaling of the fields will produce configurations with any desired energy. So the procedure in this case consists of a canonical Metropolis process, performed at a value of β chosen so as to achieve the best possible acceptance, followed by a rescaling of the fields in order to get the desired energy.

We now pass to the problem of finding the eigenvalues of the fermion matrix, which is the most CPU time consuming. This matrix is large ($V \times V$) and sparse (i.e., most of its entries are zero). The Lanczos algorithm [29]

is the best suited method for finding the eigenvalues of such matrices. Since we follow closely the version of the algorithm described in [19], we will only present its essentials in order to facilitate the implementation of our method by the interested reader.

The Lanczos diagonalization of any Hermitian matrix H takes place in two steps. First, we find the similarity transformation which tridiagonalizes the matrix; i.e., we find a unitary matrix X of size $V \times V$, such that

$$X^\dagger H X = T, \tag{36}$$

where

$$X^\dagger X = 1, \tag{37}$$

and T is a tridiagonal, real, symmetric matrix of the form

$$\begin{pmatrix} \alpha_1 & \beta_1 & & & & & \\ \beta_1 & \alpha_2 & \beta_2 & & & & \\ & \beta_2 & \alpha_3 & \beta_3 & & & \\ & & \ddots & \ddots & \ddots & & \\ & & & \beta_{V-3} & \alpha_{V-2} & \beta_{V-2} & \\ & & & & \beta_{V-2} & \alpha_{V-1} & \beta_{V-1} \\ & & & & & \beta_{V-1} & \alpha_V \end{pmatrix}. \tag{38}$$

Once the tridiagonal matrix is found, it is diagonalized with the aid of Sturm sequences. The Lanczos algorithm is based on the idea of parametrizing the unitary matrix X as a set of V column vectors, the Lanczos vectors x_i $(i = 1, ..., V)$, each of dimension $V \times 1$:

$$X = (x_1, x_2, \ldots, x_{V-1}, x_V). \tag{39}$$

The unitarity of X ensures that the column vectors x_i form an orthonormal basis that spans the V-dimensional space of the fermion matrix:

$$x_i^\dagger x_j = \delta_{ij}. \tag{40}$$

Rewriting (36) as $HX = XT$ and using (39) we obtain the Lanczos equations

$$Hx_1 = \alpha_1 x_1 + \beta_1 x_2, \tag{41}$$

$$Hx_i = \beta_{i-1} x_{i-1} + \alpha_i x_i + \beta_i x_{i+1}, \quad 2 \le i \le V-1, \tag{42}$$

$$Hx_V = \beta_{V-1} x_{V-1} + \alpha_V x_V. \tag{43}$$

These equations may be iterated V times and, using the orthonormality condition (40), lead to the following algorithm.

Step 1: Choose a starting vector x_1; the conventional choice is

$$x_1^T = (1, 0, 0, ..., 0). \tag{44}$$

Step 2: Calculate α_1, β_1, and x_2, using Eqs. (41) and (40):

$$\begin{aligned} \alpha_1 &= x_1^\dagger H x_1, \\ U_1 &:= H x_1 - \alpha_1 x_1, \\ \beta_1^2 &= U_1^\dagger U_1, \\ x_2 &= (H x_1 - \alpha_1 x_1)/\beta_1. \end{aligned} \tag{45}$$

Step 3: Now iterate step 2 $V-1$ times $(i = 2, \ldots, V)$ in order to calculate the α_i's, the β_i's and the x_{i+1}'s, using Eqs. (42) and (40):

$$\begin{aligned} \alpha_i &= x_i^\dagger H x_i, \\ U_i &:= H x_i - \beta_{i-1} x_{i-1} - \alpha_i x_i, \\ \beta_i^2 &= U_i^\dagger U_i, \\ x_{i+1} &= (H x_i - \alpha_i x_i - \beta_{i-1} x_{i-1})/\beta_i. \end{aligned} \tag{46}$$

Either sign may be used when obtaining β_i from the square root of β; conventionally we use the positive one. Also note that Eq. (43) is automatically satisfied after the last iteration (since U_V is orthogonal to all Lanczos vectors) and thus $\beta_V^2 = 0$.

Some comments are in place here. We use the Lanczos algorithm in order to find all the eigenvalues of the massless fermion Kogut-Susskind matrix $\Delta(m = 0)$ (the mass term will only shift these eigenvalues by a constant m). The matrix $\Delta(m = 0)$ is anti-Hermitian; we therefore consider the matrix $i\Delta(m = 0)$ when applying the Lanczos algorithm. We also note that $\Delta(m = 0)$ has a zero diagonal. It can immediately be seen from Eqs. (45) and (46) that this implies that all α_i's are zero. This brings about a simplification of all the above equations. Once the β's have been computed, Sturm sequences are used in order to calculate the eigenvalues of the tridiagonal matrix T.

The trouble with the Lanczos algorithm is that the rounding errors grow exponentially. The Lanczos vectors loose orthogonality after some iterations and β_V is nonzero. This implies that, after V iterations, the eigenvalues found by the algorithm are not those of the fermion matrix. One possible way out is Gram-Schmit reorthogonalization of each new Lanczos vector x_i to all the earlier ones. This solution, however, is characterized by large computer storage memory requirements and has to be ruled out.

We have adopted the alternative suggested in Refs. [30] and [19]. The technique essentially consists of iterating the algorithm W times, with $W > V$. Recall that, due to rounding errors, $\beta_V \neq 0$, and so this is always possible. Typically $W = O(2V)$. We therefore end up with a tridiagonal matrix $\tilde{T}$, which, being $W \times W$ in size, has W eigenvalues. These eigenvalues fall into three categories.

(a) The "true" eigenvalues are those which are also eigenvalues of the smaller $V \times V$ fermion matrix. They are the ones we are interested in. Since Δ is essentially almost random (its nonzero entries have been produced by a Monte Carlo simulation), it is assumed to have a nondegenerate spectrum. This assumption has always been verified *a posteriori*; no degenerate "true" eigenvalues have ever been found. Also, the eigenvalues of Δ should not depend on the number of iterations W.

(b) The "ghost" eigenvalues are degenerate eigenvalues of $\tilde{T}$ which repeat themselves several times. Their number increases with W.

(c) The "spurious" eigenvalues are eigenvalues of $\tilde{T}$, the value of which depends on W. They are recognized by the fact that they do not repeat themselves as W is increased.

So, by empirical tuning of two cutoffs, the "ghost" and

"spurious" eigenvalues are identified. The remaining set of V "true" eigenvalues are those of $i\Delta$. A property of the Kogut-Susskind fermion matrix is that the sum of the squares of its eigenvalues is $dV/4$. We use this property in order to check that all "true" eigenvalues have been found. Our check is satisfied to an accuracy of 1.5×10^{-8}. Note also that the sum of the eigenvalues to the fourth power is a linear function of the (compact) energy. In the above calculations we use 64-bit arithmetic, to ensure that we end up with all the true eigenvalues (and no others). This algorithm converges after $O(V)$ iterations; thus it has a cost of at least $O(V^2)$.

IX. PRACTICAL IMPLEMENTATION OF THE METHOD ON A TRANSPUTER NETWORK

The advantage of the method we have proposed is that supercomputers are not needed for the simulations. In fact, earlier results in the compact model have been obtained on a (shared) Vax 8650 [7, 8]. Even so, a large amount of computer time is needed: As an example, the extensive simulations presented in [9] have required the equivalent of 255 h of Cray-YMP, making unpractical their implementation on shared computers.

We therefore decided to exploit the transputer technology to build up custom made machines, partly based on commercial boards, on which to run our programs.

Recently, transputers have been widely used for a variety of applications which require parallel elaboration of data. A transputer is a chip ideally suited for multiple instruction multiple data (MIMD) architecture; it has an on-chip floating point processor and fast static internal memory, and can run at 30 MHz with 30 Mips and 2.25 MFlops peak performance. Each processor has its own private external memory; processes can run in parallel on the same transputer, or they can reside in different transputers, communicating through hardware, fast (up to 20 Mbits/s), serial channel (links). Each transputer possesses four bidirectional links, so that different communication topologies can be established. For the simulations performed, we have used different transputer boards, hosted on personal computers or on a μVax, which also manage disk I/O.

We have used the following four network configurations: (1) one single transputer board situated at the Frascati National Laboratories (LNF); (2) three 4-transputer boards situated at LNF and L' Aquila University; (3) one single transputer board connected to an 8-transputer board, situated at LNF; (4) a 64-node, 64-MBytes, electronically reconfigurable transputer-based machine (RTN) [31], situated at Zaragoza University. The collective memory of the first three configurations amounts to 49 Mbytes.

The specific implementation of our method varies according to the problem under consideration. A first possibility is by "brute force" parallelization: Either the same program is run on different transputers, with the same parameters (to increase statistics), or with different ones. This approach has partly been used in the simulations of the compact model.

The second approach used is based both on the structure of the simulations and on the topology of the boards available: It has been implemented on (two) 4-transputer networks.

The simulation consists of three logically separated parts: (a) microcanonical simulation, to produce the pure gauge configurations; (b) tridiagonalization of the fermionic matrix; (c) determination of its eigenvalues by means of Sturm sequences.

The CPU costs needed for these computations are essentially fixed (at given volume) for (b) and (c); within some limits, they are variable at will for (a).

The minimum requirement for microcanonical simulations is that configurations are well decorrelated, before diagonalization. In the simulations performed, this is always the case.

On the other hand, it turns out that, with our choice of the cutoff parameters for finding the true eigenvalues, the CPU time needed for the Lanczos tridiagonalization is approximately one-half of the time needed for the Sturm diagonalization. Moreover, the Sturm algorithm for finding eigenvalues is strictly sequential, so that even a part of the eigenvalues can be searched for independently. We have therefore divided the Sturm sequence into two parts, the first (second) searching for the first (second) half of the eigenvalues, and so the complete Lanczos process can be divided between three jobs, each running on a different transputer. The fourth transputer, which is the one linked to the host, performs the microcanonical process, to produce the next configuration, tuned so as to spend approximately the same CPU time of the other parts: In all the simulations performed, this choice has produced very well decorrelated configurations.

A disadvantage of this method is that a complete iteration is lost at the end of the simulation; on the other hand, with this subdivision it is possible to have a larger part of the code of the computer intensive parts in the on-chip memory, which allows faster execution.

In this implementation the optimal topology consists in a linear chain of four transputers, one of which is linked to the host. We have used this configuration both in compact and noncompact simulations using two 4-transputer boards at LNF.

In the noncompact case other configurations of transputers can be used, taking advantage of the fact that, since the pure gauge action is quadratic (and the measure linear) in the fields, the *same* configuration can be rescaled to have any energy required. This choice of the configurations over which to sample the distribution of the fermionic part can be shown to effectively reduce (nonphysical) fluctuations of the first derivative of the effective fermionic action. In this case the following realization has been used: A job does the microcanonical simulation to obtain the next configuration, while the previous one is distributed to the other jobs, appropriately rescaled in order to diagonalize the fermionic matrix at different, predetermined, energies. This procedure is then repeated as many times as required to have the necessary statistics. This particular realization has run both on a one + 8-transputer board at LNF and on the RTN (64 transputers) of the Zaragoza University.

X. SUMMARY AND OUTLOOK

The main objective of this paper has been to express in a systematic way the principal features of the method we have proposed. Here we summarize the most important aspects.

In our approach, the simulations with dynamical fermions have been factorized into the pure gauge generation of the configurations and the measurement of the fermionic determinant. This factorization has been accomplished by integrating over all the gauge-invariant variables which enter in the fermionic determinant, apart from the pure gauge energy. Hence configurations are generated through a microcanonical procedure, i.e., at fixed pure gauge energy. The main advantage of this method is that the most computer intensive part of the simulations does not have to be repeated for different values of coupling and flavor number. Furthermore, the above factorization allows us to have a very good decorrelation between the configurations used for the fermionic part of the computation.

The present implementation of the computation of the fermionic determinant uses the Lanczos algorithm: In this way, by finding all the eigenvalues of the femionic matrix, this part has to be performed only once for all values of the fermion mass, zero included. On the other hand, since our method is particularly well suited to the generation of decorrelated configurations, the problem of critical slowing down is under control. We stress again that, to the best of our knowledge, for observables which are not order parameters for the chiral symmetry, the zero mass value is accessible only to this method.

In practical computations the direct evaluation of the average fermionic determinant is numerically difficult, and so we resort to an approximate evaluation through an expansion in cumulants, which turns out to be an expansion in the number of flavors. It is the use of this expansion which gives us a good control over systematic errors, at least for $n_f \leq 4$. In all cases considered, the truncated expansion we use is shown to give results in very good agreement with those in the literature. In addition, while the original presentation of our method involved only the computation of thermodynamical quantities as derivatives of the free energy, an approximate evaluation of any operator not already included in the original action is possible through the use of the cumulants expansion.

The considerations we have presented for the Ising model and gauge models with a mixed action strongly suggest that the nonlocality of the original action is at least a prerequisite for the applicability of this method. We have to stress, however, that we have no description of the relation between non locality and the suppression of the fluctuations that makes the numerical evaluation possible.

Up to now this method has been used in extensive simulations of the Abelian model in $3+1$ dimensions. The advantage of its use is evident in the amount of CPU used. For the detailed study of the noncompact Abelian model we have used the equivalent of 255 h of Cray YMP. We have generated through a microcanonical procedure a total of 7771 configurations (4^4), 9370 (6^4), 6639 (8^4), 464 (10^4); these configurations have been fully diagonalized, thus allowing us to extract physical observables for *all* values of β, m, n_f. All the above configurations are very well decorrelated, the fermionic measurements being typically separated by thousands of complete microcanonical sweeps. The results presented in [9] derive from a subsample of these configurations corresponding to ~ 90 h of Cray YMP.

As already stressed, the main physical results we have obtained refer to the Abelian models in four dimensions. Work is in progress both on (2+1)-dimensional Abelian theory and on QED in $1+1$ dimensions (the Schwinger model). The latter is particularly interesting because, being solvable (at $m = 0$) in the continuum, it is ideally suited for testing the approximations introduced by the method and evaluating possible improvements.

We are also working towards the implementation of our method in SU(3). Here the main objectives are the study of the (unquenched) finite-temperature phase transition and the possible application to a full understanding of the particle spectrum beyond the quenched approximation.

As a last remark, the study of the effect of nonlocality sketched in Sec. VII shows that the present method, implemented with the cumulant expansion, has validity beyond lattice gauge theories coupled to fermions, giving, for instance, interesting results in the discussion of the Ising model beyond the mean field approximation.

ACKNOWLEDGMENTS

This work has been partly supported through a CICYT (Spain)-INFN (Italy) collaboration.

[1] M. Creutz, L. Jacobs, and C. Rebbi, Phys. Rev. Lett. **42**, 1390 (1979).
[2] D. Weingarten, in *Lattice '91*, Proceedings of the International Symposium, Tsukuba, Japan, 1990, edited by M. Fukugita *et al.* [Nucl. Phys. B (Proc. Suppl.) **26**, 126 (1992)].
[3] E. Marinari, in *Lattice '92*, Proceedings of the International Symposium, Amsterdam, The Netherlands, 1992, edited by J. Smit and P. Van Baal [Nucl. Phys. B (Proc. Suppl.) **30**, 122 (1993)].
[4] D. Weingarten, in *Lattice '88*, Proceedings of the International Symposium, Batavia, Illinois, 1988, edited by A. S. Kronfeld and P. B. Mackenzie [Nucl. Phys. B (Proc. Suppl.) **9**, 447 (1989)].
[5] E. Marinari, G. Parisi, and C. Rebbi, Nucl. Phys. **B190**, 266 (1981).
[6] D. Toussaint, in *Lattice '91* [2], p. 3.
[7] V. Azcoiti, G. Di Carlo, and A.F. Grillo, Phys. Rev. Lett.

65, 2239 (1990).
[8] V. Azcoiti, A. Cruz, G. Di Carlo, A.F. Grillo, and A. Vladikas, Phys. Rev. D **43**, 3487 (1991).
[9] V. Azcoiti, G. Di Carlo, and A.F. Grillo, Mod. Phys. Lett. A **7**, 3561 (1992); Int. J. Mod. Phys. A (to be published).
[10] V. Azcoiti, G. Di Carlo, and A.F. Grillo, Phys. Lett. B (to be published).
[11] V. Azcoiti and X.Q. Luo, in *Lattice 92* [3], p. 741. Report No. DFTUZ.92/25, 1992 (unpublished).
[12] F. Fucito, E. Marinari, G. Parisi, and C. Rebbi, Nucl. Phys. **B180**, 369 (1981); J. Polonyi and H. Wyld, Phys. Rev. Lett. **51**, 2257 (1983); G. Bhanot, U.M. Heller, and I.O. Stamatescu, Phys. Lett. **129B**, 440 (1983); P. Rossi and D. Zwanziger, Nucl. Phys. **B234**, 261 (1984); M. Grady, Phys. Rev. D **32**, 1496 (1985).
[13] S. Duane, Nucl. Phys. **B257**, 652 (1985); S. Duane and J.B. Kogut, *ibid.* **B275**, 398 (1986).
[14] S. Duane, A.D. Kennedy, B.J. Pendleton, and D. Roweth, Phys. Lett. B **195**, 216 (1987).
[15] G.G. Batrouni, G.R. Katz, A.S. Kronfeld, G.P. Lepage, B. Svetitsky, and K.G. Wilson, Phys. Rev. D **32**, 2736 (1985).
[16] Y. Oyanagi, Comput. Phys. Commun. **42**, 333 (1986).
[17] J.B. Kogut and E. Dagotto, Phys. Rev. Lett. **59**, 617 (1987); E. Dagotto and J.B. Kogut, Nucl. Phys. **B295** [FS2], 123 (1988).
[18] S.P. Booth, R.D. Kenway, and B.J. Pendleton, Phys. Lett. B **228**, 115 (1989).
[19] I.M. Barbour, N.E. Behilil, P. Gibbs, G. Schierholz, and M. Teper, in *The Recursion Method and its Applications*, edited by D. Pettifor and D. Wearie, Solid State Sciences Vol. 58 (Springer, New York, 1985), p. 149.
[20] J. Jersak, T. Neuhaus, and P.M. Zerwas, Phys. Lett. **133B**, 103 (1983).
[21] V. Azcoiti, G. Di Carlo, and A.F. Grillo, Phys. Lett. B **238**, 355 (1990); **268**, 101 (1991).
[22] L.A. Fernandez, A. Muñoz Sudupe, R. Petronzio, and A. Tarancon, Phys. Lett. B **267**, 100 (1991).
[23] M. Falconi, E. Marinari, M.L. Paciello, G. Parisi, and B. Taglienti, Phys. Lett. **108B**, 351 (1982); A.M. Ferrenberg and R. Swendsen, Phys. Rev. Lett. **61**, 2653 (1988); **63**, 1195 (1989).
[24] J.B. Kogut, E. Dagotto, and A. Kocic, Phys. Rev. Lett. **60**, 772 (1988).
[25] S.J. Hands, A. Kocic, J.B. Kogut, R.L. Renken, D.K. Sinclair, and K.C. Wang, CERN Report No. CERN-TH.6609/92, 1992 (unpublished).
[26] A. Kocic, J.B. Kogut, and K.C. Wang, Illinois University Report No. ILL-TH-92-17, 1992 (unpublished).
[27] L. Adler, Phys. Rev. D **23**, 2901 (1981); M. Creutz, *ibid.* **36**, 515 (1986); F.R. Brown and T.J. Woch, Phys. Rev. Lett. **58**, 2394 (1987); S.L. Adler, Phys. Rev. D **37**, 458 (1989).
[28] R. Petronzio and E. Vicari, Phys. Lett. B **248**, 159 (1990).
[29] C. Lanczos, J. Res. Natl. Bur. Stand. B **49**, 33 (1952).
[30] J. Cullum and R.A. Willoughby, J. Comput. Phys. **44**, 329 (1981).
[31] V. Azcoiti *et al.*, CERN Report No. CERN 92-07 353, 1992 (unpublished).